International Journal of American Linguistics

EDITOR
C. F. VOEGELIN
Indiana University

ASSOCIATE EDITORS
PAUL RIVET
Paris

HANS VOGT
Oslo

WILLIAM THALBITZER
Copenhagen

FANG KUEI LI
University of Washington

REVIEW EDITOR
HARRY HOIJER
University of California, Los Angeles

Volume 23 Number 1
January 1957

Published at the Waverly Press, Inc.
Mt. Royal and Guilford Aves.,
Baltimore 2, Md.
by Indiana University
under the auspices of
Linguistic Society of America
American Anthropological Association
with the cooperation of
Joint Committee on American Native Languages

THE INTERNATIONAL JOURNAL OF AMERICAN LINGUISTICS was founded in 1917 by Franz Boas and designed for the publication of data referring to American Indian languages. After forty years of exclusive devotion to those languages, the scope of the Journal is somewhat enlarged: to include papers on languages outside the Americas where such papers illustrate points of theory, method, and technique which have some bearing on American Indian linguistic research. The Editor will invite such contributions and will welcome correspondence from contributors who wish to publish their work in IJAL.

Primary emphasis will, however, continue to be given to data-oriented papers on American Indian languages which may be submitted, without prior correspondence, to the Editor—C. F. Voegelin, Indiana University, Bloomington, Indiana. Books and papers for review, as well as reviews themselves, should be sent to the Review Editor—Harry Hoijer, University of California, Los Angeles 24, California.

JOINT COMMITTEE ON AMERICAN NATIVE LANGUAGES

Harry Hoijer, University of California, Los Angeles (chairman; and representative of the American Council of Learned Societies)

Floyd G. Lounsbury, Yale University (representative of the American Anthropological Association)

Norman A. McQuown, University of Chicago (representative of the Linguistic Society of America)

C. F. Voegelin, Indiana University (editor of the International Journal of American Linguistics, ex officio)

The Joint Committee on American Native Languages serves as the advisory board of the International Journal of American Linguistics and is a body for planning, coordinating, and encouraging research and publications relating to the languages native to the Western Hemisphere.

International Journal of American Linguistics is published quarterly at the Waverly Press, Inc., Mt. Royal and Guilford Avenues, Baltimore 2, Md., by Indiana University. Subscriptions ($5.00 a year) will be received by the Waverly Press, Inc., Mt. Royal and Guilford Avenues, Baltimore 2, Md., or by the Business Manager, Mrs. Elsie Dosch, Indiana University, Bloomington, Ind.

Entered as second class matter May 11, 1945 at the post office at Baltimore, Maryland under the Act of March 3, 1879

International Journal of American Linguistics

VOLUME XXIII January, 1957 NUMBER 1

A TENTATIVE TYPOLOGY OF ETYMOLOGICAL STUDIES

YAKOV MALKIEL

UNIVERSITY OF CALIFORNIA

0. Introduction
1. Classification by scope
2. Classification by material
3. Classification by degree of complexity
4. Etymology and general linguistics
5. Bibliography

0.1. Not only is etymology one of the oldest linguistic disciplines, but some kind of lay curiosity about word origins seems to have everywhere preceded organized research by a margin of millennia. Even if one disregards as scientifically irrelevant the magic, mythical, and allegoric preoccupations with etyma, the remaining corpus of researches is forbiddingly vast. A typology of these studies that would attempt to do equal justice to all such legitimate genres, fashionable slants, and personal styles as are on record can hardly be undertaken without years of preliminary canvassing. The unpretentious sketch here presented, which embodies the outgrowth of a few weeks' thinking and searching, at best aspires to provide a minimum of tentative orientation. It bears thorough revision, with an aim, above all, at more widely spread and more evenly distributed documentation.

0.2. The multitude of possible fruitful approaches (only these have I thought worth-while taking here into account) has been broken down, by way of an experiment, according to three major criteria: first, the varying scope of the chosen inquiries; second, the kind and amount of material adduced as evidence; third, the inherent degree of transparency of each problem at issue. The order of these three criteria may be freely inverted. Overlapping, I hope, if not entirely eliminated, has been severely reduced. Whether still other systems of typological classification may be successfully introduced into etymology (and, if so, how many and which), I do not undertake to determine at this point. For a while I toyed with the idea of arranging typologically the varying conditions of actual etymological discoveries (chance findings, results of systematic surveys, testing of combinations of sounds, etc.), but abandoned it upon realizing that most meritorious workers, for the sake of decorum, have been satisfied with presenting carefully retouched versions of their original drafts, without expatiating on the intuitional or anecdotal side of their labors.

0.3. The classification here essayed refers strictly to approaches, not to solutions. Etyma that come up for mention are quoted exclusively for the sake of the technique involved in identifying them. Unless the opposite is stated, the bases do not bear any stamp of approval or acceptance, beyond a tacit acknowledgement of respectability.

0.4. For clarity's sake, a border-line, however thin, must be traced between the domains of etymology proper and of lexicology (the latter, in turn, most scholars make it a point to distinguish from lexicography, the art of compiling dictionaries). Etymology has been traditionally defined as the elucidation of word origins, by means (its modern practitioners would add) of several increasingly rigorous, mutually complementary controls. By virtue of its dependence on phonemic and morphemic transformations, etymological analysis partakes of historical grammar. To the extent to which it heeds the evidence of word histories ('lexical biographies, trajectories'), etymology partially coincides with diachronic lexicology. More specifically, etymological research is coterminous with the inquiry into the in-

cipient phase of a given word history, a phase ordinarily shielded from direct observation ('lexical embryology'). On this assumption, a lexical study (including one historically slanted) that fails to concern itself explicitly with the crystallization of a lexical unit or with its initial penetration into a new speech medium would fall outside the scope of etymology. In contrast, an etymological study is invariably a contribution to lexicology, except that occasionally the full examination of the record may be dispensed with or curtailed (e.g., if the complete material has already been assembled in a generally accessible publication, or if a school of thought prefers to separate the core of the linguistic analysis from the philological or anthropological shell in which the selected forms were found embedded).

1.0. Historically the oldest and, until this day, the typical and most widely consulted etymological analysis is the entry found in a representative etymological dictionary. Monolingual dictionaries emphasizing etymology (e.g., S. de Covarrubias's Tesoro for Spanish, 1611; G. Ménage's Origines for French, 1650 and 1694) precede by a wide margin multilingual, i.e., comparative, dictionaries, such as F. Bopp's, for Indo-European (1847) and F. Diez's, for Romance (1853).[1] Midway between these extremes is the dictionary essentially devoted to a single language, but conveying, apropos of congeners, stray bits of information on cognate languages, such as W. W. Skeat's work for English (1882, 1910), from which etymological data may be gleaned on German, Dutch, and Scandinavian, or the revised editions of W. Gesenius's Hebräisches und aramäisches Handwörterbuch (e.g., 16th ed. of 1915), with its profusion of collateral forms drawn from other Semitic languages.

1.01. The average entry in an etymological reference work may be couched in popular language and still represent the distillation of genuine scholarly research, as is true of F. Kluge's highly readable dictionary for German (including the editions revised by A. Götze), of O. Bloch's and A. Dauzat's, for French (1932 and 1938, respectively, with later revisions in both instances), and of B. Migliorini's, for Italian (1950, 1953). Or the phrasing, though erudite, may be unencumbered by bibliography, a level of specificness chosen by A. Ernout and A. Meillet for Latin, particularly in the first two editions of their Dictionnaire étymologique (1932, 1939; the 3d ed., of 1951, refers the reader less sparingly to learned publications). There are other possibilities: each entry may contain a plain selective bibliography, as in A. Prati's Italian dictionary (1951), or a bibliography reflecting the author's avowed bias, as in the Indogermanisches etymologisches Wörterbuch (1949 ff.) of J. Pokorny (who admits his prejudice against studies operating with laryngeals); or else it may be burdened by a practically exhaustive bibliography, a probably not inaccurate assessment of the apparatus in J. B. Hofmann's recently completed revision of A. Walde's Latin dictionary.

1.02. A comparable measure of freedom prevails with respect to the selection of words subjected to scrutiny. The selection may range from very broad to quite narrow, except that the former possibility, on the whole, is typical of pioneer works (e.g., Diez's) and the alternative, of their modern counterparts. Further limitation of the stock is possible through explicit or implied concentration on certain ancestral layers (cf. A. de Cihac's segregation of Latin and non-Latin [Hungarian, Turkish, Albanian, etc.] strains in Rumanian, which S. Puşcariu, in 1905, strove to emulate in a dictionary left incomplete; or R. Dozy's and W. H.

[1] Strictly speaking, no historical, let alone etymological, dictionary can be truly monolingual, since, by definition, several time levels are involved. For the sake of convenience, one may attach this unprecise label to such etymological dictionaries as illustrate the convergence of different lexical currents upon a single literary language (say, of Ibero-Basque, Latin, Gothic, Western Arabic, and Amerindian on Spanish).

Engelmann's joint venture in the Hispano-Arabic field), through emphasis on arbitrarily chosen grammatical categories (thus, the torso of R. J. Cuervo's Spanish syntactic dictionary consistently favors verbs, pronouns, and prepositions, never failing to etymologize them), or through confinement to a single regional or social dialect (e.g., G. Rohlfs's Calabrian vocabulary and his dictionary of Modern Greek speech islets in southern Italy).[2]

1.03. Each word chosen for succinct etymological identification may be presented in a separate entry. In languages favoring derivation and composition this procedure, entailing awkward repetition, is characteristic of reference works in which etymological information is distinctly subordinate to other services, e.g., of Webster's New International Dictionary (2d ed., 1950), of the Spanish Academy Dictionary (18th ed., 1956), of A. Darmesteter, A. Hatzfeld, and A. Thomas's Dictionnaire général (1890–900), and of comparably planned standard works. The alternative for the straight etymologist is to group words by lexical families, of various degrees of inclusiveness, a course charted by the overwhelming majority of contemporary workers in the Indo-European field. Through use of cross-references some scholars even go one step farther and discuss under a master entry semantically related words of diverse provenience, e.g., names of the weekdays, of the cardinal points, certain mutually conditioned sets of pronouns, and the like. J. Corominas has probably advanced too far along this road of concentration (1954 ff.).

1.04. The entry may vary in size, starting from an epigrammatic remark and swelling to monographic proportions, as in W. von Wartburg's monumental thesaurus (1928 ff.); and within the entry, the space reserved for the etymological discussion may be relatively exiguous, as in the laconic comments of the Thesaurus Linguae Latinae, or excessive, as in Corominas's discursively phrased dictionary. It also happens that rather detailed and fairly sketchy entries (and the corresponding discussions) alternate, according to the objective intricacy of each problem or to the author's private preferences; E. Gamillscheg's and again J. Corominas's dictionaries are examples in point.

1.05. In a monolingual etymological dictionary the base may function as the entry, a rarely exploited possibility (W. von Wartburg's dictionary, which at first glance seems to illustrate this architecture, in fact subsumes a wealth of dialect forms under each entry, hence can hardly rate as monolingual); or the product acts as the entry, a standard practice for several centuries. If a work is bipartite and bidirectional, both arrangements may be tried out consecutively, cf. the unique pattern which V. García de Diego adopted in 1954. In multilingual etymological dictionaries dealing with a single language family, one expects to find the material listed by inferred prototypes or by the recorded archetypes of an ancestral language; a combination of these two categories is also possible, cf. W. Meyer-Lübke's Romanisches etymologisches Wörterbuch (1911–20, 1930–5), whereas G. Gröber's earlier pilot project, restricted to Vulgar Latin reconstructions, is reminiscent, despite its miniature size, of, say, Pokorny's Indo-European dictionary. The pioneers sometimes proceeded differently: Bopp's Glossarium Sanscritum (in later editions, titled more graphically Glossarium comparativum linguae Sanscritae) consistently started out from Sanskrit words, subordinating to them in rank their I.-E. cognates, a hierarchy no longer recognized. Diez, less understandably, followed suit by recording pan-Romanic words under the corresponding Italian entries. Whichever the sequence preferred, an appended alphabetical index, if

[2] Infrequently, selective etymologizing serves the opposite purpose of eliminating categories of words deemed irrelevant. This policy M. R. Haas has adopted on a minor scale in her Tunica Dictionary, especially as regards borrowings from American-French.

competently executed, ordinarily enables the reader to visualize the connections in reverse order.

1.1.0. Given the chronological precedence of the dictionary over the lexical monograph, the etymological note, as a genre of scholarly output, originally served the subsidiary purpose of either filling gaps in the latest authoritative dictionary or of proposing conjectures at variance with those advocated by recognized lexicographers. Thus, N. Caix, as a Romanicist, and R. Thurneysen, as a Celticist, were fundamentally content with enriching, elaborating, and critically sifting the information furnished by Diez's dictionary (1853, 1861, 1869–70), to which A. Scheler, an etymologist in his own right known from his research in French, independently provided a formal comprehensive Supplement. C. Salvioni, intimately acquainted with North Italian and Raeto-Romance dialects, and O. Nobiling, at home in more than one sense with (Brazilian-) Portuguese, extended a similar treatment to G. Körting's less meritorious compilation (1891, 1901, 1907). On the Spanish side alone, A. Castro, V. García de Diego (1923), and R. Menéndez Pidal (1920) went each with a fine comb over the original edition (1911–20) of Meyer-Lübke's comparative dictionary. Even more assiduously, M. L. Wagner re-analyzed hundreds of entries pertaining to the Sardinian lexicon in the original as well as in the revised (3d) edition of that same dictionary.

1.1.1. In preference to supplementary notes some etymologists have poured out a great wealth of fresh or revived hypotheses (either impressionistically stated or supported by evidence) in extended book reviews of etymological dictionaries. Among Romance scholars, known for their sympathy with this practice, the technique reached its height some thirty to forty years ago. The crowning accomplishments were possibly the critiques of J. Jud, saturated with felicitous ideas and trenchant formulations. Conversely, the writings of J. Brüch and L. Spitzer slanted in this direction rate, at best, as controversial. A lone American counterpart was the pioneer M. A. Todd's review article.

1.1.2. Another manifestation of this vogue of small etymological gleanings correlated to a major dictionary project has been the pre-publication, as it were, of a discovery. The etymological note was launched as a trial balloon, frequently years ahead of the corresponding entry in a planned dictionary—if that dictionary came at all into existence. C. Michaëlis de Vasconcelos herself, by the colorful titles which she gave to some collections of her notes, characterized them as so many preliminaries to a future Luso-Hispanic dictionary, which she eventually no longer found the strength to write. E. Gamillscheg also had recourse to this preview technique, eliciting from his colleagues criticism from which, in his definitive dictionary (1926–8), he sometimes refused to benefit. Corominas's scattered notes, after the lapse of a few years, logically became the ingredients of his currently published dictionary.

1.2.0. Although crudely manufactured historical grammars did exist long before 1800, the flowering of this particular projection of linguistic facts falls into the period 1850–1915. Since historical phonology (and, to a lesser extent, morphology) may be described as a skillfully arranged mosaic of etymological equations, a few of them upheld or impugned in extended comments, a manual of historical grammar so viewed contains a well-sampled collection of, by and large, irreducibly short etymological notes. (These notes occupy a lower level of abstraction than the respective schemes of sound shifts and of inflectional, derivational, compositional, and syntactic changes.) Some distinguished historical grammars have actually been used over the decades as mines of etymological information, formulaic and denuded of all frills. In a book structured like R. G. Kent's Old Persian, which wraps together, in a single handy package, gram-

mar, inscriptional texts, and lexicon, it is even difficult to decide whether the section on phonology or the section on the lexicon contains the greater wealth of etymological nuggets.

1.2.1. What matters most in such context is the weighty etymological decision, i.e., the support of one out of two or several rival etymological hypotheses which an authoritative grammarian is willing to throw into the balance. A grammar may also contain a few original lexical solutions; yet, all told, it is unusual for an experienced and versatile scholar to avail himself of a grammatically slanted work as the ideal or the exclusive medium for propounding etymological conjectures. True, exceptions are on record, including É. Benveniste's Origines from which Indo-Europeanists, I understand, have culled numerous fresh explanations. Equally rare and, in the long view, uneconomical is the circumstantial treatment of grammatical points in an essentially etymological study, a procedure exemplified by the excursuses in J. Gilliéron's Genealogy of the Gallo-Romance names of the bee and by the very core of H. M. Flasdieck's more recent book-length report on the British street-name Pall-Mall. Here the announced project threatens to serve as a mere pretext or façade for an utterly different inquiry.

1.3.0. Given the early and operationally close ties between textual criticism, historical grammar, and etymology (all three occasionally subsumed, and not by laymen alone, under the vague label 'philology'), the careful edition of texts has yielded, as a frequent by-product, a harvest of etymological observations. The optimal conditions prevail in the case of ancient languages known almost entirely from a single monument or from a severely limited number of texts.

1.3.1. One can envisage two typical situations. The first involves an etymological glossary attached to a text or to the collected writings of an author—an arrangement at present out of fashion, since it normally entails casual presentation of etymologies, as in W. Förster's vocabulary to Chrétien de Troyes (an exception, as regards the degree of elaboration, is Menéndez Pidal's masterly vocabulary to the Cantar de Mio Cid). For a while, such etymological glossaries were commonly appended to chrestomathies and textbooks, as an aid to beginners, and enjoyed added prestige if the compiler had earned a reputation as an independent worker (C. Voretzsch, J. D. M. Ford). The more recent tendency, a wholesome innovation, has been to disregard etymology completely in such word lists and to concentrate on the exact synchronic delineation of meanings, constructions, and paradigms (L. Foulet, C. C. Marden)—a noteworthy parallel to other modern trends favoring minute description in preference to hasty reconstruction.

1.3.2. The second situation involves dubious individual words, on which the disputed dating, localization, and authorship of a text may hinge. Thus, generations of Slavic scholars (at present, especially R. Jakobson and, among his associates, K. H. Menges) have busied themselves with the etymologies of some intrinsically recalcitrant or insecurely transmitted words peculiar to the Old Russian Igor' Tale. It is impossible to separate the etymological inquiry into Old French mun(t)joie (war-cry or religious exclamation) from the painstaking textual interpretation of the Chanson de Roland, as shown by the recent studies of L. H. Loomis and of her critics.

1.3.3. If exacting textual criticism demands again and again the etymologist's craft and equipment, neatly edited texts, by way of compensation, provide him with effective tools for his own digging. The bridge between Gr. makár(ios) *blessed, happy*, on the one hand, and Old Port. magar, Old Span. maguer(a) *although*, on the other hand, may appear impassable, but the ingenuous speech of a mountaineer girl, recorded (1034d) and stylized by that keen observer Juan Ruiz (ca. 1330), plus our collateral knowledge of conservative Italian

usage (magari) provide the otherwise unavailable intermediate stages.

1.4. The autonomous etymological note, neither designed to supplement, correct, or announce a dictionary, nor forming part of, or preparatory to, a historical grammar or a glossary, but written for its own sake, arose ca. 1850 and reached its all-time peak a few decades later. Among Romanicists, K. A. F. Mahn may be the earliest devotee of this genre; A. Thomas in France, H. Schuchardt in Austria, and C. Salvioni in Italy were its unsurpassed masters, who created three distinctive styles, and J. M. Piel in Portugal beside J. Haust in Belgium have been its most seasoned rearguard strategists. The etymological note purports to solve the problem of a word origin with a minimum display of heavy erudition; its elegance is inherent in its terseness. No attempt is made to piece together the full history of the word; only characteristic glimpses are given (unfortunately, some less than accomplished scholars confuse 'characteristic' with 'selected at random, so as to favor a solution arrived at a priori'). Almost invariably, the note carries with it a margin of risk; if, at its best, it solves a problem with an admirable economy of means, it remains, at its weakest, a flashy rather than witty piece of writing which, once refuted, evaporates without leaving a trace, since the array of facts it boasted, in all likelihood, was chosen haphazardly, to fit a single whim. The ultimate decay of this genre, as the lexicon of each language becomes more open to historical inspection, is inevitable, insofar as only an insignificant part of the total of word histories may be compressed into a few sentences without loss of vital information. The not too numerous words that lend themselves to this approach are either those exhibiting no major change of meaning, no external influence of associative processes, scarcely any effect of linguistic diffusion (and their etymologies were established long ago in the better-known languages), or words of anecdotal background like macadam and sandwich.

1.5.0. The modern substitute for obsolete sequences of loosely strung etymological notes has been planned research in etymology. Such research, aside from the writing of a dictionary as a life-long project, may assume different forms. The simplest to describe is the full-sized etymological article devoted to the origin and early development of a single word, in particular an article calling for plentiful documentation and likely, as a result, to swell to monographic proportions.[3] E. J. Dobson's attempt to place Engl. boy *churl, servant,* (after 1400) *male child* in an appropriate social setting, to delimit its fluctuating semantic ambit, and to determine its overtones at successive time levels, with a view to tracing it cogently to Anglo-Norm. *(em)buié, -boié *slave, serf,* from buie < Lat. boia *fetter, chain,* illustrates this isolative approach at its most conscientious.

1.5.1. Ordinarily it is the fragmentary or otherwise aberrant record of a word that cautions the linguist to pause and grant it the benefit of microscopic examination, splitting his attention between the slightest oscillations of form, meaning, and area. Thus, the volume of material which N. Törnqvist amassed in an attempt to analyse Germ. Ufer *shore,* orig. *slope,* as a blurred Germanic compound (< ô- + fê²r) rather than a unitary cognate of Gr. 'ēpeiros *continent,* derives its justification from the disquieting absence of the word from Old High German. M. B. Emeneau's etymological curiosity about Proto-Drav. *va- *to come* and *ta- *to give, bring* is due to their singular morphological anomaly, attributable to the loss of the so-called transition suffixes *-a- and *-ar- (except in Kui) and to the ensuing mixture of variants according to a formula changing from language to language. The

[3] A separate article may also enter into a specific relation to a dictionary project; thus, W. von Wartburg's study of the Gallo-Romance progeny of Lat. albus *white* was astutely timed so as to show its readers the advantage of the novel classificatory method which was to preside over the author's current spadework on a new-style dictionary.

painstaking establishment of an etymology may serve an ulterior purpose, as when M. R. Haas undertakes her meticulous study of the Proto-Gulf word for *water* preëminently in defense of her own classification of Gulf languages (Muskogean, plus Natchez, Tunica, Chitimacha, and Atakapa), a classification significantly divergent from J. R. Swanton's, M. Swadesh's, and E. Sapir's.

1.5.2. More widely accepted and far more economical in terms of the long-range investment of human resources is the publication of a cluster of closely related notes, especially those based on first-hand knowledge of the same background, past or present, on the acquaintance with the same group (not necessarily family) of languages, and on the application of the same analysis to the chosen corpus of data (field notes, excerpts from texts preserved, gleanings from dialect maps, etc.). These characteristics apply, at two extremes of a spectrum of tastes, to H. Rheinfelder's inquiry into the Italo-French lay religious vocabulary and phraseology rooted in the Catholic liturgy and to B. E. Vidos's and H. and R. Kahane's researches in the spread of Mediterranean nautical terms (involving probings, aside from Southern Romance, also into Middle and Modern Greek, Anatolian Turkish, Vulgar Arabic, and the lingua franca of seaports). V. Bertoldi's bold studies in the archaic Mediterranean substratum, T. Frings's contributions (in response to Jud's programmatic article of 1914) to Latin-Germanic lexical contacts, and A. Steiger's explorations along the axis Middle and Near East–Western Europe signal the high-water mark of pertinent European specialization. Two of the leading American representatives of this approach have been D. S. Blondheim and W. A. Read, each endowed with an unusual combination of skills: the former, in A. Darmesteter's wake, brought to bear his knowledge of Latin, Old French, and Biblical beside Late Hebrew on the cryptic rabbinical glosses; the latter, without staking out any sweeping claims, in fact launched comparative Caribbean lexicology by setting in motion, from the colonial Mississippi Delta as his operational base, the disparate resources of Louisiana (and Canadian) French, Louisiana Spanish, West Indies Dutch, Haitian Creole, Cherokee, Island Carib, etc.

1.5.3. Refinement in a different direction may be achieved when such etymologically rebellious words as are suspected of interlocking or echoing one another are attacked in pairs or triads or when a whole lexical family is jointly subjected to stringent, yet imaginative analysis. Thus, Jud, in a brilliantly executed (if not wholly convincing) article, recognized the dual anomaly in the coinage of It. menzogna/Fr. mensonge vs. Sp. Port. mentira *lie*, two deviations from the norm which he was inclined to attribute to a single pattern, twice activated, of analogical interference. The unknowns here were not the root words, but the derivational schemes. Ernout, with an etymologist's flair for detecting and bringing into full view the hidden and the inconspicuous, restored the full semantic range, blurred at the surface, of Lat. uenus *erotic desire, sexual intercourse*, Venus *goddess of love* by focusing not only on the nuclear formation, but on the satellites as well: uenustus *gracious*, uenerius *pertaining to physical love*, uenārī *to strive for* > *to hunt*, uenēnum *poison*, uenia *favor, pardon*, etc. Studies in lexical polarization, some of them etymologically motivated, fit quite smoothly into this context.

1.5.4. Apart from folk etymology, homonymy is presumably the most formidable natural enemy of the etymologist. One way of facing its inroads squarely is to select, as the avowed specific object of an etymological experiment, the unraveling of a homonymic knot or tangle. This has been done by many scholars on a variety of scales: tentatively by F. Lecoy and almost casually by J. Corominas with respect to Hisp. cach-, yet quite elaborately, long before them, by E. Richter in regard to Romance bur(d).

1.5.5. Between the two World Wars, European continental scholars engaged suc-

cessfully in the genetic analysis of synonyms and near-synonyms as a focus of lexical research, with special reference to dialect geography. For a while this 'onomasiologic' approach, of which B. Quadri recently submitted a mid-way appraisal, was contrasted with semasiology, i.e., linguistic (as opposed to philosophical) semantics; conceivably synonymics would have been a less pretentious and more apposite designation. The typical starting point was the map of a linguistic atlas, which, under favorable conditions, offered up to twenty carefully localized and transcribed lexical types, some of them etymologically transparent, others opaque. These latter a trained worker etymologized with the aid of the complicated machinery available for each language; eventually a broad attempt was made to reconstruct the actual sequence of events, in terms of linguistic geology (or paleontology). In this kind of research project, etymology represented a vital link, but solving an equation was scarcely the ultimate goal of each venture. What mattered principally in the end was the emergence of a clear-cut, harmonious lexico-cultural pattern, with special attention paid to the coördinates of time and space. These studies, characteristically titled 'The names (or designations) of . . .', were for the most part extended doctoral dissertations, often bearing the hallmark of a formal academic exercise. A few won wide recognition, including E. Eggenschwiler's in the Romance names of the *bat*.[4] An early isolated counterpart in this country is H. R. Kahane's genetic inventory of the names of the *cheek* (and, indirectly, of the *jaw*) in Italian dialects, with marginal inclusion of Greek and Germanic bases. The subtlest interpretation of etymologically challenging maps we owe to K. Jaberg, who also excels in the combined study of lexicology and folk beliefs (witness his forthcoming article on the birthmark). Exploratory onomasiologic advances beyond the confines of a single language family (ordinarily Romance or Germanic) have been few by comparison; one such daring experiment is C. Tagliavini's inquiry, cutting across Indo-European, Semitic, and North African languages, into the names of the pupil of the eye.

1.5.6. In its most ambitious form, the etymological monograph follows the trajectory of a single word and its satellites across formidable language barriers and over a period of centuries and even millennia. The fortune of this genre is inevitably tied up with the etymologist's grasp of cultural and linguistic diffusion. If more than trivial lacunae remain to be filled in, the resultant itinerary may best be symbolized by a dotted line, as when É. Benveniste traces Class. Lat. obrussa, Low Latin obryza *cupellation of gold*, lit. *earthen container*, through Greek to Hittite-Hurri ḫubrušḫi, from ḫabur-, ḫubur- *earth*, via a lost Semitic link; independently, the possibility of Hittite-Greek borrowings also aroused E. Sapir's keen curiosity, though he was even less explicit about the channel. B. Hasselrot's inquiry (a classic of its kind) into the migratory fruit name which ultimately yielded Engl. apricot leads the reader from Imperial Rome via Greece and the Near East (at first Hellenistic, later Arabicized) along the North African coastline and across the Straits of Gibraltar to Spain, and from there, through the Pyrenees, to Paris as a new focal point of pan-European expansion. The infiltration of indigenous New World words such as Carib, cannibal, and Shakespeare's Caliban, into the various European languages throughout the colonial period was the object of P. Henríquez Ureña's stimulating and sensitively phrased reports. Two book-length projects, A. S. C. Ross's intricate reconstruction of the Eurasiatic trajectory of *ginger* and H. M. Flasdieck's

[4] The rash of these monographs, coinciding with the crest of the vogue of linguistic atlases, seems to be almost over. In the Romance domain the latest that has come to my attention is H. Kröll's on the designations of *drunkenness* in colloquial Portuguese (understandably, an investigation placing heavier stress on metaphors than on etymologies).

simultaneous pan-European study of *tin* and *zinc* (which German pairs off more effectively: Zinn and Zink) represent an extreme degree of sophistication along this line. A more modest American equivalent would be E. W. Bulatkin's recent scrutiny of the Romance words for *nuance*, a lexical study broadening out, despite its sharp focus, into the history of the Graeco-Romance manufacture of colors and dyes.

1.5.7. Like every genre, the etymological article carries with it certain drawbacks and advantages. Compared with the independent note, it tends to displease the linguistic analyst by the fact that the nuclear problem of word origins dwindles to a disproportionately small size, becoming the terminal point of a long and sinuous preliminary inquiry: the straight, quasi-mathematical analysis appears dangerously diluted. The advantages outweigh this deficiency: while the intuitional conjecture, supported by scant (if any) evidence, is as easily deflated by adverse criticism as a punctured balloon, the profusely documented etymological monograph is resistant, if not entirely immune, to such attacks. Granted that the analysis of the foundation may be marred by an error, the superstructure need not crumble at the first stroke of a pickaxe wielded by a hostile hand (cf. the unequal value of the chapters in A. Farinelli's account of Sp. marrano *crypto-Jew*). The etymological article, like any other full-scale biography, rarely derives its only value from a definite solution of the problem of parentage.

1.5.8. The chief advantage of the article over the corresponding dictionary entry is the incomparably greater measure of freedom that the linguist enjoys in shaping it. Each case receives individual attention: the word is embedded in a cultural context uniquely its own, the volume of documentation is determined by the degree of regularity in the changes undergone, etc. Characteristically, most advances in etymological technique have lately been achieved in such medium-scale ventures inviting experimentation and tactical maneuvering rather than in the rigid framework of a dictionary, whose author is necessarily haunted by the restraining ideal of consistency.

2.0. Another criterion for classifying etymologies is the quality and quantity of the material utilized. In this respect, pioneer researchers were content with a scientifically valid, but somewhat rudimentary approach, subjecting forms extracted from ancient texts to increasingly rigorous historical and comparative analyses (vertical and horizontal axes). The progress at that (one is inclined to say, philological) stage was measured: a) by the expansion of the corpus of texts newly discovered or for the first time critically edited, with the concomitant revision of their vocabularies; b) by the refinement of textual criticism, especially in dealing with infrequent or erratic forms apt to provide fresh clues to etymologies; c) by a shift of emphasis in the selection of canvassed texts, away from standard genres favoring trite words and in the direction of technical treatises (bestiaries, lapidaries, recipes, etc.) which contain a store of words defying the most fastidious etymologist.

2.1. As Indo-Europeanists started ranging beyond the realm of extinct languages and as modern (especially, Romance) philology broke away from classical studies, dialect material began to gain recognition in etymological debate, initially only at its fringes. A later generation of scholars contrived a state of equilibrium between philological and dialectological documentation, by setting out to fill the lacunae, unavoidable in lexical transmission, through systematic exploitation of present-day conservative dialects. This approach presupposes the elevation of dialect study, once (and, in part, until this day) the domain of well-meaning amateurs, to the rank of a respectable scholarly endeavor, a change by and large effected by G. I. Ascoli for southwestern Europe and by his peers elsewhere. The new

combined approach was championed, at the turn of the century, by a brilliant and cosmopolitan galaxy of Romanicists (R. J. Cuervo, R. Menéndez Pidal, J. Leite de Vasconcelos, A. Thomas, W. Meyer-Lübke, and many others, including their immediate associates and followers). H. Schuchardt, over a period of sixty feverish years, not only extracted every imaginable benefit from this pool of resources, juggling with a virtuoso's deftness his Romance, Slavic, Turkish, Hungarian, and Creole data, but boldly applied the method to the social and regional stratification of Latin. Contemporary front-line Latinists like Ernout, Hofmann, Niedermann, and the Swedes have studded their inquiries into the records of Antiquity with precise references to medieval Latin and vernacular formations and, as a second choice, to modern lexical material, ranging from Romance to Basque, Vulgar Arabic, and Berber, to the extent to which Latin etymology (in particular, the study of submerged words) may profit from such data. In the Indo-Aryan field P. Tedesco practices a similarly balanced approach, striving to cast floods of light on Sanskrit words with the aid of New Indic regionalisms.

2.2. The Experimental Age in etymology began when scholars, familiar with these two traditional methods and skilled in the use of their mutually corroborative evidence, deliberately set out to confine themselves to a single technique, either entirely new or novel in its isolation. The gamut of driving forces behind this voluntary abandonment of comfortably secure ground included the quest for fresh insights from which incalculably valuable results might be expected; the lingering fear that the most satisfactory technique, if employed beyond the point of saturation, may eventually become sterile; and, probably, some workers' special delight in ever renewed contact with slices of material particularly pleasing to them (archival data, accurately tagged specimens of material civilization, records of native speech elicited from informants, bibliographic apparatus of a well-stocked research library, etc.).

2.3. J. Gilliéron was less of a revolutionary than his shocked contemporaries were led to assume (much of his colorful phrasing he owed to É. Littré, homonymics before him had been cultivated by F. A. Coelho, and the cartographic presentation of linguistic facts is an invention usually credited to G. Wenker). However, the abrupt decision to dispense altogether with philological evidence (including dialect glossaries) in favor of unretouched field notes projected onto a series of maps as the sole guide to reconstruction was Gilliéron's private whim. He rarely attacked etymologically obscure formations, preferring to focus attention on the chronological sequence of words paired as synonyms or homonyms, real or potential, exact or approximate. But the method intrinsically lent itself to extension toward purely etymological analysis.

2.4. Gilliéron's researches produced sharper repercussions among Romance linguists than in other quarters. Here the work of an entire team of dynamic and talented etymologists (K. Jaberg, M. L. Wagner, J. Jud, W. von Wartburg, E. Gamillscheg, G. Rohlfs) was aimed at the reconciliation of cartographically streamlined dialect geography with the subtler and more complex methods of conventional etymology.

2.5.0. A different experiment consisted in the self-imposed limitation to special types of historical records, deemed more dependable than others. This applies to accurately dated and localized notarial documents, preserved in the original, which some younger etymologists, after 1920, began using to the strict exclusion of literary texts, which they held to be either too late for the chosen purpose or linguistically blurred by the uninhibited interference of successive copyists. Collections of charters had been compiled before by philologists sensitive to paleographic niceties (e.g., by P. Meyer in 1909, covering sections of southern France). E.

Staaff, two years earlier, cogently drew grammatical conclusions from his meticulous edition of well-chosen Old Leonese texts. Menéndez Pidal succeeded in basing, by and large, on these inconspicuous texts (some of which he himself had made accessible as early as 1919) a massive grammatical edifice of Proto-Spanish dialects (1926, 1929, 1951), with stray lexico-etymological appendages. At this point bold experiments with this selective method began all over Europe, not each of them, in retrospect, commanding an equal share of interest. G. Tilander wisely applied it to Old Port. rousar and relatives *to rape*, a verb coined in chanceries and apparently unknown in other environments. He was less well advised in relying on it when engrossed by the crux Old Span. es-, tres-, tras-quilar *to shear*: here the inclusion of texts other than municipal ordinances not only would have been legitimate, but might have considerably improved our perspective on a word etymologically elusive, but by no means limited to a single social sphere.

2.5.1. P. Aebischer and J. Hubschmid, both from Switzerland, operate almost invariably with this sampling of material. The former has made himself known preëminently, the latter, exclusively, as an etymologist. These two scholars, instead of spreading their curiosity over the entire lexical stock of a given language, select small or middle-sized word groups, tracing with astonishing patience each formation through volume after volume of almost unreadably monotonous cartularies and inventories. At the end, each interprets the emerging pattern diachronically, as a seasoned dialect geographer would do a linguistic map. Although their material is similar, their practices of selection and their territorial and chronological preferences happen to clash.

Aebischer concentrates on words of Latin ancestry, moves with ease through the maze of Low Latin records (favoring, as he should, those tinged by Romance speech habits), keeps within his purview texts written in the medieval vernaculars, and adds a mere sprinkle of modern dialect words. His tacitly assumed goal is extra-precise stratigraphy based on minute observation, not necessarily clarification of nebulous word origins, so that he only incidentally stoops to the discovery of new etyma. He seems to earmark words for investigation on the grounds of their repeated occurrence in his favorite medium (notarial texts) rather than of any inherent affinity of a higher order. His conclusions, though not invulnerable, appear on the whole sensible, conservative, discerning, and delicately shaded.

Hubschmid distills, from his vast readings, the small residue of those notoriously troublesome words that seem traceable to a pre-Latin, preferably even a pre-I.E., substratum in western Europe and attempts to link them, somewhat recklessly, to geographically very distant cognates (or putative cognates). The aim almost looks like some kind of acrobatic performance; the result is etymological research with an unprecedented number of unknowns, and a minimum of controls; the punishment, a disturbingly high incidence of (in part, admitted, in part, not easily verifiable) errors; the atonement: an exceptionally rich collection of neatly extracted raw data, useful even to those fellow-workers who may disagree with the underlying scale of values, in general, and with the specific findings announced, in particular.

2.6. Another experimental limitation has been attempted in regard to onomastic, especially toponymic, data. Circumstances may impose this lopsided emphasis, as when a word of the ancestral language has survived only in proper names. One such very tentative tabulation of Latin petrifacts G. Sachs prepared years ago for the Iberian peninsula. In substratum studies this restriction occasionally yields gratifying results, cf. Menéndez Pidal's note on Proto-Basque Javier-Chabarri. In most instances, however, it is desirable and practicable to blend onomastic with dialectal shreds of

evidence; witness Menéndez Pidal's and Hubschmid's independent analyses of Sp. cueto *protuberance*, divergent in all respects except in the only possible way of sampling the data. Toponyms, freely admitted into articles, have been slow in penetrating the barrier of etymological dictionaries; M. Vasmer's, for Russian, marks a promising departure from this inertia.

2.7. Etymological diagnosis based on the geographic factor alone has not, to my knowledge, yet been essayed. But geographic conditions may play the decisive role in assigning a debated word to a given lexical stratum, a classification which, in turn, may entail the definitive choice of a suitable base. Geographic proximity to another language, for instance, may tip the scales in favor of an assumption of borrowing: an obscure French dialect word generously represented in northeastern patois but consistently absent from the west and the south is more than likely to be of Frankish descent. More problematic has been the attempt of the Italian neolinguists to infer patterns of relative chronology (conducive, in the last analysis, to etymological hypotheses) from the contrasting areal configuration of present-day rival words.

2.8.0. Patterns of meaning and imagery as a clue to etymology have so strongly obsessed the minds of some scholars as to have allowed them to ride roughshod over other, not one whit less important, considerations. The ever-present conflict between semantic and phonic plausibility in assessing the merits of a conjecture was pointed up, half a century ago, by a memorable duel between A. Thomas, resolved to elude the temptation of semantic analogy, and H. Schuchardt, cheerfully prepared to yield to the lure. The skirmish ended in an impasse; meanwhile, scholars have begun realizing that the semantic tool, before being tried out again, needed a good deal of refurbishing. Only recently (1951) Benveniste reminded us, among other hazards, of the inadequacy, in etymological discussion, of such terms as 'abstract' and 'concrete', of the dangerous discrepancies between lexical contours in the language explored and in the explorer's language used as a frame of reference, of the oft-neglected, indissoluble ties between meaning and construction.

2.8.1. If the linguistic logician instinctively shrinks from this path, the poetically endowed etymologist returns to it again and again. In his ability to solve etymologies or, at least, to suggest plausible rapprochements through the re-creation of faded imagery J. Trier probably has no living peer; witness his powerful treatment of the etymological motif *mother–children* ~ *root-stock* (or *stem*)–*sprigs*, with special reference to Lat. māter, mātrīx, māteriēs. This is visual linguistics at its most attractive. Unfortunately, Trier, despite his training, allows himself to be carried away by his fertile imagination. The inclusion of matta within the māt- complex is hazardous (as he himself senses), and even the bare mention of matula in this context seems preposterous. The poetic evocation of patterns of imagery simply lacks those controls that have helped make the acoustic branch of linguistics a possibly less pleasing, but, beyond doubt, more effectively organized science. The appeal of this method to the dormant graphic sensitivity in the linguist is healthy, and in conjunction with other, less intuitive approaches it acts as a precious stimulant; but it has, one must admit, wrought havoc with not a few linguistic facts when applied in isolation.

2.9.0. To solve an etymological problem in abeyance, a worker may, finally, decide on collecting, with equal zeal, linguistic and correlated non-linguistic facts. The slogans Wörter und Sachen, Wort- und Sachatlas have centered attention about material civilization as a clue to etymology. Within the broad boundaries of this method one may substitute anatomical and even kinship terms for the names of tools and containers: what matters is not the fabric, shape, purpose, or palpability of the object, but the planned acquisition, on the part of the ex-

plorer, of technical knowledge beyond the confines of linguistics, and his ability to bring it to bear on lexical reconstruction. The worker's attention need not be focused on the terminal point of the inquiry, the actual etymological commitment: F. Krüger, in his Leonese and Pyrenean word studies, for the most part stops short of this final step, while Wagner, in cataloguing the Portuguese names (and corresponding varieties) of the mill-hopper, presses the etymological issues with genuine alacrity. Etymology so slanted easily becomes a semilinguistic discipline, the remaining half being claimed by a potentially unlimited number of complementary specializations. V. Bertoldi, to write convincingly on ancient and modern phytonyms, became an expert botanist, pharmacologist, and culture historian all in one; P. Barbier's protracted studies in European fish names prompted him to qualify as an amateur ichthyologist; É. Benveniste, in defining, with his usual precision, Upper Yukon animal names, showed respectable knowledge of Alaskan wildlife; O. J. Tallgren-Tuulio, before examining Hispano-Arabic names of stars and constellations (for the sake of their transcription), deepened his understanding of ancient Graeco-Oriental astronomy. It is difficult to imagine the combination of a long-established science and of linguistics that an etymologist could not, under attractive conditions, be induced to straddle.

2.9.1. In one border-case the etymologist, aside from contributing to the identification of a specific word origin, finds himself delving into the history and theory of linguistics as a whole, the matrix of his own discipline. This happens when the critical digest of earlier scholarly pronouncements, normally a prefatory piece ancillary to the actual analysis of the chosen linguistic specimen, assumes, by way of exception, such proportions as to outweigh, in richness of results, the principal operation. Far and away the most rewarding section of Menéndez Pidal's inconclusive article on the toponym Madrid is his penetrating criticism of untenable Greek, Latin, Germanic, and Arabic conjectures; one can award a similar prize to Vincent's study of the equally controversial name Gaul. The literature on Fr. aller and congeners *to go* became so prolific and scattered half a century ago (it has doubled in volume since then) that G. Stucke had to write an interim report for those in need of quick but thorough orientation. Primary linguistic material here almost yields in interest to (if one may expand on L. Bloomfield's dictum, without injecting into it his animus) secondary responses to etymology.

3.0. Possibly the boldest way of hierarchizing etymological problems is to arrange them in the order of decreasing transparency. (The most challenging cases are indubitably those known as obscure, but the most rewarding, on the strength of the actual yield, are probably found at a respectable distance from either extreme.) Equations like Rum. deget < Lat. digitu *finger* and Rum. piept < Lat. pectu(s) *breast, chest* present no complications whatever, furnishing choice material for the establishment of sound correspondences along the time axis; but they fail to whet the very tools with which the etymologist must handle more resistant slices of material. At the opposite end of the scale are the notorious cruxes, i.e., problems still unsolved and, so far as foreseeable, often doomed to remain insoluble (cf. Ernout and Meillet's pessimistic comment on Lat. adulārī *to fawn upon*, aestu-, aesti-māre *to appraise, value*, car-cer, -car *jail*, imitārī *to portray, act like*, etc.). The intermediate categories are those most in need of careful analysis.

3.1. Alongside the truly simple solution one finds the deceptively simple solution, usually because some unevenness which might have served as a clue has been smoothed out—through the agency of folk etymology or of another associative process. Thus, Lat. aboriginēs *original inhabitants*, āmanuēnsis *clerk*, and pedeplāna *rooms on*

the ground-floor are no longer believed to reflect ab orīgine, ā manū, pede plānō, in this order; Russ. samoyéd, on the surface meaning *one who eats himself*, is actually a disguise for something utterly different, etc. In this sector, sound correspondences lose most of their usefulness as guides to word origins, the historical record remaining the all-important clue.

3.2. Nothing teases an etymologist more than facing a partially transparent situation. The Latin fish-name aci-pēnser, -pensis has been plausibly explained as including a classifiable head and an unclassifiable tail. Fr. lou-ange *praise*, mélange *mixture*, and mensonge *lie* are inseparable, by common-sense standards, from the corresponding verbs louer, mêler, mentir (genetically, all of them perfectly diaphanous), but their derivation has been a stumbling-block.[5] In Romance, minor disturbances of this kind are frequently caused by void non-final suffixes (interfixes), difficult of explanation, cf. Fr. fort-er-esse (orig. -ece), Sp. fort-al-eza *strength, fortress*.

3.3. The hindrance to clear-cut analysis need not be an anomalous derivational or compositional pattern. In the case of Fr. laisser, It. la-sciare, -ssare, Old Port. leixar, Old Span. dexar, Rum. lăsa, etc. *to let, leave* < Lat. laxāre *to slacken* beside Old Span. dexar (> dejar), Port. Cat. deixar, Gasc. dechá, Sard. dassare, Cal. Sic. dassari *to let, leave*, the initial d- of the smaller second series has taxed the ingenuity of etymologists since 1600. It took scholars considerable time to pinpoint the difficulty: leaving out the skeptics, agnostics, and spineless compromisers, one may distinguish between those dissociating genetically lexar and dexar, tracing the former to laxāre and the latter to dēserere, dēsinere (Covarrubias), *dēsitāre (Diez), *dē-iectiāre, -iexāre (Ricc), or *dēcessāre (Nicholson), a school of thought now practically extinct; and those who jointly link lexar and dexar to laxāre, on five different, in part mutually exclusive assumptions: 1) broad identification (Sarmiento, Schuchardt 1868, Coelho, Tailhan . . .); 2) sound change l- > d- (a. tendential pan-Romanic shift: Reinhardstoettner; b. consonant dissimilation: Cornu, Menéndez Pidal 1900; c. Sabinian idiosyncrasy: Silva Neto; d. consonant assimilation: Corominas); 3) dī-, *dē-laxāre as the sole source of dexar (Ascoli, Guarnerio, Gorra . . .); 4) blend of laxāre and *dēlaxāre (Schuchardt 1891, Ford 1900, Menéndez Pidal 1908 . . . Wagner); 5) contamination of laxāre by an unrelated word (a. dē: Ford 1911; b. dare: Meyer-Lübke, Spitzer; c. Arab. mudağğan: Schürr). This loose league of scholars, divided on many details of their platform, at least, has essentially agreed to hem in tightly the anomaly of l- > d- in this one word, a considerable methodological progress.

3.4. From here the number of difficulties in a given word may indefinitely increase. Thus, the consensus of most scholars is that Fr. soin *care* and be-soin *need* share a morpheme, but the origin of soin as well as of be-, viewed separately, and the compositional scheme of besoin remains dubious, a wide margin of three unknowns left uneliminated.

4. This typological sketch, however modest, represents an attempt to bring about a rapprochement between etymology and general linguistics. Etymologists are frequently recruited from the ranks of scholars inherently wary of generalization and become imbued with such knowledge as does not lend itself smoothly to abstraction. In contrast, modern linguistics has been moving at a swift pace in the direction of increasingly abstract formulation. Hence two conflicting scales of values and not a few conflicts of personalities. But a complete break is undesirable, because etymology divorced from general linguistics degenerates easily

[5] Sp. Port. mentira *lie*, in its unique relation to the verb mentir, presents a parallel aberration. For an attempt to remove the obstacle to neat classification see Romance Philology **6**. 121–72 (1952–3).

into a genteel game or into a discipline ancillary to history, while linguistics deprived of its etymological roots tends to assume a degree of abstractness which the study of languages in its full richness hardly invites. There may thus be some point to our seeking to discover recurrent types, i.e., constants, in etymology, even if we agree to characterize it as the domain of individuality in language history.

5. BIBLIOGRAPHY[6]

Aebischer, P. *Estudios de toponimia y lexicografía románica* (Barcelona, 1948).

Barbier, P. Noms de poissons. Notes étymologiques et lexicographiques, *Revue des langues romanes* **51.** 385–406 (1908), **52.** 97–129 (1909), **53.** 26–57 (1910), **54.** 149–90 (1911), **56.** 172–247 (1913), **57.** 295–342 (1914), **58.** 270–329 (1915), **63.** 1–68 (1925), **65.** 1–52 (1927), **67.** 275–372 (1935).

Benveniste, É. *Origines de la formation des noms en indo-européen*, 2 vols. (Paris, 1935–49).

—— Problèmes sémantiques de la reconstruction, *Word* **10.** 251–64 (1954).

—— Le terme 'obryza' et la métallurgie de l'or, *Revue de philologie*[3] **27.** 122–6 (1953).

—— Le vocabulaire de la vie animale chez les Indiens du Haut Yukon (Alaska), *Bulletin de la Société de linguistique de Paris* **49:** 1. 79–106 (1953).

Bertoldi, V. Contatti e conflitti di lingue nell'antico Mediterraneo, *Zeitschrift für romanische Philologie* **57.** 137–69 (1937).

Brüch, J. Bemerkungen zum etymologischen Wörterbuch E. Gamillschegs, *Zeitschrift für französische Sprache und Literatur* **49.** 290–318 (1926), **50.** 299–355 (1927), **52.** 393–483 (1929).

Bulatkin, E. W. The Spanish word *matiz*: its origin and semantic evolution in the technical vocabulary of medieval painters, *Traditio* **10.** 459–527 (1954).

Caix, N. *Studi di etimologia italiana e romanza; osservazioni ed aggiunte al vocabolario etimologico delle lingue romanze di F. Diez* (Florence, 1878).

Castro, A. Adiciones hispánicas al Diccionario etimológico di W. Meyer-Lübke, *Revista de filología española* **5.** 21–42 (1918), **6.** 337–45 (1919).

Cihac, A. *Dictionnaire d'étymologie daco-romane*. Vol. 1: Éléments latins; Vol. 2: Éléments slaves, magyars, turcs, grecs-moderne et albanais (Frankfurt ª/M., 1870–9).

Corominas, J. Indianoromanica, *Revista de filología hispánica* **6.** 1–35, 139–75, 209–54 (1944) [on *cach-*: pp. 33–4].

—— Problemas del diccionario etimológico, *Romance Philology* **1.** 23–38, 79–104 (1947–8).

Cuervo, R. J. *Disquisiciones sobre filología castellana*, ed. R. Torres Quintero (Bogotá, 1950).

Darmesteter, A., and D. S. Blondheim, *Les gloses françaises dans les commentaires talmudiques de Raschi*, 2 vols. (Paris, 1929–Baltimore, 1937).

Dobson, E. J. The etymology and meaning of *boy*, *Medium Aevum* **9.** 121–54 (1940).

Dozy, R., and W. H. Engelmann, *Glossaire des mots espagnols et portugais dérivés de l'arabe*, 2nd rev. ed. (Leyden, 1869) [1st ed. of 1861 by Engelmann alone].

Eggenschwiler, E. *Die Namen der Fledermaus auf dem französischen und italienischen Sprachgebiet* (Bern, 1934) [for Hispanic addenda see *Hispanic Review* **19.** 238–63, 323–40, of the year 1951].

Emeneau, M. B. The Dravidian verbs 'come' and 'give', *Language* **21.** 184–213 (1945).

Ernout, A. 'Venus, venia, cupīdō', *Revue de philologie*[3] **30.** 7–27 (1956).

Farinelli, A. 'Marrano', *Studi . . . Pio Rajna* 491–555 (Florence, 1911); *Marrano: storia di un vituperio* (Geneva, 1925) [for criticism see *Journal of the American Oriental Society* **68.** 175–84, of the year 1948, and B. Migliorini, *Lingua Nostra* **10.** 50, of the year 1949].

Flasdieck, H. M. 'Pall Mall': Beiträge zur Etymologie und Quantitätstheorie, *Anglia* **72.** 129–383 (1954).

—— *Zinn und Zink; Studien zur abendländischen Wortgeschichte* (Tübingen, 1952).

Foerster, W. *Kristian von Troyes; Wörterbuch zu seinen sämtlichen Werken* (Halle ª/S., 1914; rev. ed., by H. Breuer, 1933).

Ford, J. D. M., ed. *Old Spanish readings, elected on the basis of critically edited texts* (Boston, 1911, 1939) [Etymological Vocabulary, pp. 177–312].

Foulet, L., in J. Bédier, ed. *La Chanson de Roland*, Commentaires (Paris, 1927) [Glossary, pp. 321–501].

Frings, T. Germania Romana. *Teuthonista*, Suppl. 4 (Halleª/S., 1932).

Gamillscheg, E. Französische Etymologien (I–II), *Zeitschrift für romanische Philologie* **40.** 129–90, 513–42 (1919–20); **41.** 503–37, 631–47 (1921–2).

García de Diego, V. *Contribución al diccionario hispánico etimológico* (Madrid, 1923; reprinted 1943).

Gilliéron, J. (with M. Roques). *Études de géo-*

[6] All easily identifiable standard reference works have been omitted.

graphie linguistique d'après l'Atlas linguistique de la France (Paris, 1912).

—— *Généalogie des mots qui désignent l'abeille* (Paris, 1918).

Gröber, G. Vulgärlateinische Substrate romanischer Wörter, *Archiv für lateinische Lexikographie und Grammatik*, Vols. 1–6 (1884–9).

Haas, M. R. The Proto-Gulf word for 'water' (with notes on Siouan-Yuchi), *International Journal of American Linguistics* **17.** 71–9 (1951).

—— Tunica dictionary, *University of California Publications in Linguistics* **6:** 2. 175–332 (1953).

Hasselrot, B. L'abricot. Essai de monographie onomasiologique et sémantique, *Studia Neophilologica* **13.** 45–79, 226–52 (1940–1).

Haust, J. *Étymologies wallonnes et françaises* (Liége, 1923).

Henríquez Ureña, P. *Para la historia de los indigenismos* (Buenos Aires, 1938) ['Caribe': pp. 95–102].

Hubschmid, J. Studien zur iberoromanischen Wortgeschichte und Ortsnamenkunde, *Boletim de Filologia* [Lisbon] **12.** 117–56 (1951).

Jaberg, K. *Aspects géographiques du langage* (Paris, 1936).

—— The birthmark in folk belief, language, literature, and fashion, *Romance Philology* **10:** 4 (1956–7).

Jakobson, R. L'authenticité du 'Slovo', *Annuaire de l'Institut de philologie et d'histoire orientales et slaves* **8.** 235–60 (1945–7) [esp. pp. 240–62].

Jud, J. It[alien] 'menzogna'; fr[ançais] 'mensonge'; esp[agnol] 'mentira', *Vox Romanica* **11.** 101–24 (1950).

—— Probleme der altromanischen Sprachgeographie, *Zeitschrift für romanische Philologie* **38.** 1–75 (1914–7).

—— Review of W. Meyer-Lübke, *Romanisches etymologisches Wörterbuch*. *Archiv für das Studium der neueren Sprachen* **127.** 416–38 (1911).

Kahane, H. R. Designations of the cheek in the Italian dialects, *Language* **17.** 212–22 (1941).

Kahane, Renée. Italienische Marinewörter im Neugriechischen, *Archivum Romanicum* **22.** 510–82 (1938) [in criticism of D. C. Hesseling, 1903].

Kahane, H. and R. The Mediterranean term SURGERE 'to anchor', *Romance Philology* **4.** 195–215 (1950–1) [p. 195: bibliography of earlier publications].

Kent, R. G. *Old Persian: grammar, texts, lexicon* (New Haven, 1950).

Kröll, H. Designações portuguesas para 'embriaguez' (Coimbra, 1955) [reprinted from *Revista portuguesa de filologia* **5.** 27–85, **6.** 74–134, **7.** 17–120].

Lecoy, F. Étymologies espagnoles, *Romania* **68.** 1–17 (1944–5) [on *cach-*: pp. 1–8].

Loomis, L. H. The oriflamme of France and the war-cry 'monjoie' in the twelfth century, *Studies . . . for Belle de Costa Greene* 67–82 (Princeton, 1954); cf. eadem, *The Romanic Review* **41.** 241–60 (1950) and *Speculum* **25.** 437–56 (1950); W. A. Nitze, *Romance Philology* **9.** 11–7 (1955–6); J. Harris, ibid. **10: 3** (1956–7).

Mahn, K. A. F. *Etymologische Untersuchungen auf dem Gebiete der romanischen Sprachen* (Berlin, 1855).

Marden, C. C., ed. *'Libro de Apolonio', an Old Spanish poem*, Part 2, Elliott Monographs 11–2 (1922) [Vocabulary: pp. 67–185].

Menéndez Pidal, R. La etimología de 'Madrid' y la antigua Carpetania, *Revista de la Biblioteca, Archivo y Museo* **14.** 3–23 (1945); reprinted and revised in *Toponimia prerrománica hispana* 189–220 (Madrid, 1952).

—— Etimologías españolas, *Romania* **29.** 334–79 (1900).

—— 'Javier-Chabarri', dos dialectos ibéricos, *Emerita* **16.** 1–13 (1948); reprinted in *Toponimia prerrománica hispana* 233–50.

—— Notas para el léxico románico, *Revista de filología española* **7.** 1–36 (1920).

Menges, K. H. *The oriental elements in the vocabulary of the oldest Russian epos, the 'Igor' Tale'*. Preface by R. Jakobson (New York, 1951).

Niedermann, M. Ueber einige Quellen unserer Kenntnis des späten Vulgärlateinischen, *Neue Jahrbücher für das klassische Altertum* **29.** 313–42 (1912).

Nobiling, O. Berichtigungen und Zusätze zum portugiesischen Teil von Körtings 'Lateinisch-romanischem Wörterbuch', *Archiv für das Studium der neueren Sprachen* **124.** 332–45 (1910), **125.** 154–7, 393–7 (1910), **126.** 175–9, 424–32 (1911), **127.** 181–8, 371–7 (1911–2).

Piel, J. M. *Miscelânea de etimologia portuguesa e galega*, 1st series (Coimbra, 1953).

Quadri, B. *Aufgaben und Methoden der onomasiologischen Forschung*. Romanica Helvetica 37 (Berne, 1952).

Read, W. A. *Louisiana-French* (Baton Rouge, 1931).

Rheinfelder, H. *Kultsprache und Profansprache in den romanischen Ländern* (Geneva, 1933).

Richter, E. Die Bedeutungsgeschichte der romanischen Wortsippe 'bur(d)', *Sitz.-ber. Vienna* **156:** 5. 1–138 (1908).

Ross, A. S. C. *'Ginger', a loan-word study* (Oxford, 1952).

Sachs, G. Restos latinos en nombres de lugares españoles, *Archivum Romanicum* **19.** 437–40 (1935).

Salvioni, C. Postille italiane al vocabolario latino-romanzo, *Memorie dell' Istituto Lombardo* **20.** 255–78 (1899).

Sapir, E. Greek 'atýzomai', a Hittite loan word, and its relatives, *Language* **12.** 175–80 (1936).

Scheler, A. Supplement to Diez, *Etymologisches Wörterbuch* . . . , 4th ed. (Bonn, 1878) and 5th ed. (1887).

Schuchardt, H. Romanische Etymologien I–II, *Sitz.-ber. Vienna* **138:** 1. 1–82 (1897), **141:** 3. 1–222 (1899).

Spitzer, L. Aus Anlass von Gamillschegs 'Französischen Etymologien', *Zeitschrift für romanische Philologie* **42.** 5–34 (1922).

Steiger, A. Aufmarschstrassen des morgenländischen Sprachgutes, *Vox Romanica* **10.** 1–62 (1949).

Stucke, G. *Französisch 'aller' und seine romanischen Verwandten; eine kritisch-etymologische Untersuchung.* Diss. Heidelberg (Darmstadt, 1902).

Tagliavini, C. Di alcune denominazioni della 'pupilla'. Studio onomasiologico, con speciale riguardo alle lingue camito-semitiche e negro-africane, *Annali dell'Istituto universitario orientale di Napoli* (N. Ser.) **3.** 341–78 (1949).

Tallgren [-Tuulio], O. J. Los nombres árabes de las estrellas y la transcripción alfonsina, *Homenaje a Menéndez Pidal* **2.** 633–718 (Madrid, 1925).

Tedesco, P. Sanskrit *muṇḍa-* 'shaven', *Journal of the American Oriental Society* **65.** 82–98 (1945).

Thomas, A. *Mélanges d'étymologie française* (Paris, 1902); 2d rev. ed. (1927).

Thurneysen, R. *Keltoromanisches; die keltischen Etymologien im 'Etymologischen Wörterbuch der romanischen Sprachen' von F. Diez* (Halle ª/S., 1884).

Tilander, G. Origen y evolución del verbo 'esquilar', *Studia Neophilologica* **9.** 48–65 (1936–7).

—— Rausar, rousar, rouxar, roixar; rauso, rouso, rojo, roxo, roiso, *Boletim de Filologia* [Lisbon], **6.** 188–97 (1939–40).

Todd, H. A. [William A.] Knapp's Spanish etymologies, *Modern Language Notes* **1.** 117–20, 142–6 (1886).

Törnqvist, N. Zur Etymologie von n[eu]h[och] d[eutsch] 'Ufer', *Studia Neophilologica* **13.** 253–90 (1940–1).

Trier, J. *Lehm: Etymologien zum Fachwerk* (Marburg, 1951).

—— *Holz: Etymologien aus dem Niederwald* (Münster, 1952) [pp. 136–43: 'māter'].

Vasconcelos, C. Michaëlis de. Contribuïções para o futuro dicionário etimológico das línguas hispânicas, *Revista Lusitana* **11.** 1–62 (1908).

Vasconcelos, J. Leite de. Notas filológicas, *Revue hispanique* **4.** 209–14 (1897), **5.** 417–29 (1898).

Viana, A. d. R. Gonçalves. *Apostilas aos dicionários portugueses*, 2 vols. (Lisbon, 1906).

Vidos, B. E. *Storia delle parole marinaresche italiane passate in francese* (Florence, 1939).

Vincent, A. 'Gallia' et 'Gaule', *Revue belge de philologie et d'histoire* **27.** 712–26 (1949).

Voretzsch, C. *Einführung in das Studium der altfranzösischen Sprache*, 6th ed. (Halle ª/S., 1932); Glossary: pp. 372–411 [1st ed., 1901].

Wagner, M. L. Rettifiche ed aggiunte alla terza edizione del REW . . . , *Archivum Romanicum* **19.** 1–30 (1935); **20.** 343–58 (1936); **24.** 11–67 (1940).

—— Sobre os nomes da 'moega' nas línguas iberorománicas, *Biblos* **24.** 247–65 (1948).

Wartburg, W. von. 'Albus' und seine Familie in Frankreich, *Zeitschrift für romanische Philologie* **41.** 182–92 (1921–2).

YUMAN LANGUAGES I: FIRST IMPRESSIONS

WERNER WINTER
UNIVERSITY OF KANSAS

0. This is the first in a series of articles dealing with linguistic data obtained during the summer of 1956 on a fieldtrip to Arizona and California, which was made possible through a generous grant from the University of Kansas and the hospitality of the Museum of Northern Arizona, Flagstaff, Arizona.[1]

The purpose of the trip was to establish contact with speakers of the various Yuman languages in the area;[2] to collect basic linguistic data, thereby in part duplicating and—with the help of modern equipment and the advantage of a shorter time lapse between the recordings of the related idioms—rechecking data published by Kroeber, Spier, and others;[3] to secure, wherever possible, more extensive texts for later evaluation and publication.

Within the time limits of two and one-half months, it was possible to collect materials from speakers of the following idioms[4] (in brackets the number of informants worked with): Havasupai (3), Walapai (7), Yavapai (5), Mohave (2), Maricopa (2), Yuma (1), Cocopa (2), Diegueño (5), 'Campo' (5). The choice of rather to increase the number of informants from each idiom than to concentrate on one or two particularly able persons for more intensive work was necessitated by the survey character of the whole project: only in this way did it seem possible to avoid a generalization of idiolectic features in the speech of the preferred informant; even a quick test with three other persons showed, e.g., that the leveling of /s/ and /š/ done by an otherwise very good 'Campo' informant at Viejas, California (absent only in a few forms like ʔəšmá. *I sleep*, as contrasted with ʔsî *I drink*) was strictly idiolectic: the other informants clearly distinguished between nɪmšáp *white* and nɪmsáp *dawn* (Walapai nʸɪmsáv : nʸɪmθáv).[5] The obvious disadvantage of being prevented from supplementing the extensive approach by intensive work on details could be fully eliminated only in the case of Yavapai and Walapai, where excellent informants could be used over some greater length of time.

The texts obtained vary greatly in character: ranging from simple narration of

[1] Thanks are due to C. F. Voegelin for proposing the project and for numerous suggestions during the summer's discussions at Flagstaff; to Jerry Hopkins for making accessible to me part of his Havasupai materials; to the Reverends Carillo and Nardi for facilitating contacts with speakers of Diegueño and 'Campo' and for their deep personal interest; to the great number of Indians whose patience and cooperation made possible the success of the work undertaken.

[2] For data on the languages in question and references to prior literature see A. L. Kroeber, Classification of the Yuman languages, UCPL 1.3. 21–40 (Berkeley-Los Angeles, 1943).

[3] Note especially: Kroeber, *op. cit.*; L. Spier, Comparative vocabularies and parallel texts in two Yuman languages of Arizona, Univ. of New Mexico Publications in Anthropology, nr. 2 (Albuquerque, 1946); in addition the various ethnographic monographs on individual tribes, which contain valuable linguistic information. A. M. Halpern's study of Yuma (IJAL 12–13. 1946–47) has been freely used; the very fact of its being available, however, was the reason why only a few hours of fieldwork were spent on this language, which in turn led to virtual exclusion of Yuma from the discussions in this first article.

[4] I use the term 'idiom' where an identification as to 'language' or 'dialect' would either be premature or inconvenient, as in statements covering both individual languages and dialects.

[5] The transcription in this first article is in broad phonetics, neither trying to mark all features of the allophones occurring, nor attempting a phonemic analysis yet. Wherever such an analysis seems appropriate, though, it has been performed and duly marked.

events in the life of the informant to expert oratory and formal rendering of tribal lore and legends, they should furnish important insights into questions of narrative techniques and levels of style. Two sets of texts seem especially interesting, though for almost opposite reasons: both in Maricopa and in Yuma, the informants provided texts that were translations—in the one case rendering in Yuma an account on tribal traditions, which had been composed in English, in the other translating a passage from the Bible; in the latter instance, the translation was not formulated spontaneously, but had been written down by the informant, and was now read off. In contrast to these materials of a highly acculturated character, the very substantial body of text recorded in Walapai gives the impression of complete genuineness; very remarkable seems the fact that not only the originals appear well narrated, but that also the translations furnished by the informant are well-told stories in their own right.

1. Linguistic information sought for, by necessity mostly pertained to the lower levels of analysis. Thus, while the phonological data secured would seem sufficient to warrant statements on at least segmental phonemics of most individual languages and consequently to enable us to reconstruct, with a reasonable degree of certainty, a considerable number of proto-forms, morphological data seem just to suffice to analyze texts in the better documented idioms, but not yet to venture comprehensive statements in comparison and reconstruction. As a result, the comments in this first article will deal almost exclusively with phonological and lexical matters.

2. With the exception of the California group, which Kroeber himself tended to consider as possibly not forming a unit, the classification established by Kroeber seems well borne out by the more diversified materials obtained during my fieldtrip. As tentative evidence on the lexical level the following data are of interest:

Among 100 lexical items chosen at random, though with considerable emphasis on words from Swadesh lists, the following number were identical in two idioms respectively (identity in this context is understood to mean complete formal convertibility of two corresponding items; no mere similarity or apparent relatedness is admitted as establishing identity).[6] See Chart on page 20.

The minimum number of correspondences, 50, is found between Diegueño and Walapai, and also with other languages of the entire group the correspondences found in Diegueño remain extremely low (it is to be mentioned here that 37 of the 100 items are shared by all individual languages), so that Diegueño is well set off as a language of its own, not forming a subgroup with any other idiom listed here. The only possible tie would be with 'Campo' (see below), with which Diegueño shares 77 items; since, however, 'Campo' is equally closely connected with Cocopa, with which Diegueño shares only a low number of items, 'Campo' and Diegueño cannot be combined into a subgroup; rather is 'Campo' to be considered as being the connecting link between Diegueño and Cocopa. Cocopa in turn is clearly set off against the languages to the north of it, though the number of agreements with Maricopa is fairly high; here again, an intermediate position between two languages or groups of languages is suggested by the balance between the correspondences of Cocopa with Maricopa on the one hand and with 'Campo' on the other, with only a low number of correspondences between 'Campo' and Maricopa. Maricopa itself is closely connected with Mohave, without this language having a high number of agree-

[6] In a few instances, the informant's response to a particular question consisted of two items, each with a different cross-group affiliation. In such cases, both items were counted; the figures arrived at are, therefore, not exact percentages.

	Havasupai	Walapai	Yavapai	Mohave	Maricopa	Cocopa	Diegueño
Walapai	95						
Yavapai	92	91					
Mohave	63	63	62				
Maricopa	58	57	57	85			
Cocopa	56	54	54	65	76		
Diegueño	51	50	51	52	52	61	
'Campo'	59	57	57	56	64	77	77

ments with Cocopa; we are thus safe in locating a major boundary between Mohave and Maricopa on the one side and Cocopa on the other, which corresponds exactly to Kroeber's setting up a River group (Mohave, Maricopa, Yuma) separate from the Delta group (Cocopa, Kahwan). The number of correspondences between Mohave-Maricopa and the northern idioms is rather low, scarcely above the figures for even the most distant languages of the group; this indicates another clearcut boundary for the Mohave-Maricopa group. The northern idioms among themselves are more closely in agreement than any other pair of languages under discussion; this strongly suggests the possibility that in the case of Havasupai-Walapai-Yavapai we are not dealing with three separate languages, but rather with diaects of one and the same language.

3. Of Kroeber's California group, no language but Diegueño could be studied. In spite of all efforts, no Akwa'ala could be located north of the Mexican border. (One aged person purported to be an 'akwȧł matxá. *northern A.* was of no help; as far as could be determined, her speech showed only slight deviations from the local (Viejas) idiom, and these deviations did not point in the direction of Akwa'ala as depicted in Gifford's publication, but rather toward Yuma or Maricopa.) Even so, the Yuman-speaking area in Southern California (disregarding here entirely Yuma proper) turned out to be linguistically diversified: while materials obtained from the Diegueños of Santa Ysabel, San Felipe, and Iñaja were uniform, those acquired from informants from Viejas, Los Conejos (now living at Viejas), and Campo (now living at Jamul) proved to be different from the northern Diegueño data, but consistent among themselves. Even though, in terms of tribal affiliation, these people south of the Diegueños proper seem to be considered Diegueños, too, linguistically they are better taken as a separate entity, as has been pointed out in the preceding paragraphs. While it may be that this idiom is identical with Kroeber's 'Kamia', this identification is not absolutely certain, and until a survey of the territory just south of their habitat can be made, and possible further affiliations can be established, a tentative designation of this idiom as 'Campo' (after one of the reservations where it is spoken) may be in order.[7] That 'Campo' is not a dialect of Diegueño, cannot be conclusively established by the test of mutual intelligibility or its absence: while one Diegueño informant claimed to understand the speech of a particular 'Campo' who had been mentioned in our conversation, his wife, who is a Diegueño like himself, disclaimed such understanding, and her husband clearly confirmed this statement. As we have seen above, lexical data rather indicate for 'Campo' a position intermediate between Diegueño and Cocopa than a closer connection with either. Closer examination of the data dealt with above

[7] According to Steve Ponchetti, one of my Diegueño informants, Diegueño proper is spoken at Barona Ranch, Capitan Grande, Iñaja, Santa Ysabel, San Felipe, Mesa Grande, and San Pasqual, 'Campo' at Los Conejos, Viejas, Campo, Jamul, Manzanita, Cuyapaipe, and Sequan.

yields the following detailed results: Of 102 words from the lexicon of 'Campo' 59 are shared with both Diegueño and Cocopa, 18 with Diegueño alone, 18 with Cocopa alone, 7 with neither (of which one item in turn is shared by Diegueño and Cocopa against all other languages). Of the 18 items shared with Diegueño alone, 6 are further shared with at least one member of the northern group, and 3 more with languages from the north and central areas, which leaves 9 items shared with Diegueño only. Of the 18 items shared with Cocopa alone, one is further shared with the north, 6 with north and centre, and 7 with centre only; this leaves 4 items shared with Cocopa only. Of the 7 items shared neither with Diegueño nor with Cocopa, 2 are shared with the northern group, leaving thus 5 items attested only for 'Campo'. This gives us a total of 18 items not found outside the languages of the south and west and therefore suspected to be innovations originating here. Keeping in mind the tentative character of the sampling and the analysis, we seem entitled to the conclusion that a slightly greater number of what seem to be at least in part innovations were shared by Diegueño and 'Campo' than by 'Campo' and Cocopa, but that the number shared between the latter two combined with the number of apparent innovations found so far to exist in 'Campo' alone, prevents us from connecting 'Campo' any closer with Diegueño than as a neighboring related language.

What is thus suggested by lexical data, is further supported by first results of phonological analysis:

Diegueño is set off against all other Yuman languages recorded by showing x in positions where the other idioms have š (Maricopa, Yuma, Cocopa) or s (Mohave, Yavapai, Walapai, Havusupai).[8] This characteristic feature is not shared by 'Campo', which, in this respect, goes rather with Cocopa (and Maricopa-Yuma). On the other hand, no merger has taken place of this š with the sound corresponding to MoYaWH *θ*, MaYu s, as in Cocopa; rather does 'Campo' here go with Diegueño (and Maricopa-Yuma) in preserving the contrast.

Another unique Diegueño feature not shared by 'Campo', is the occurrence of morpheme-final -lʸ corresponding to glottal stop in other languages. (This feature of Diegueño appears to be restricted to positions where the related languages allow a reconstruction of a sequence Vy before the final glottal stop; relevant examples are: D ʔɪlʸ *wood*: M oʔaʔí·, Y aʔɪʔíʔ; D ʔuwílʸ *stone*: M oʔavíˑ, Y aʔuwíʔ; D ʔɪsílʸ *salt*: M oʔaθʔíˑ, H ʔɪθíʔi; D sènʸ xamélʸx *young woman*: W(h)méʔɛ *boy*. The only instance where Diegueño has lʸ in a position following what has to be reconstructed as V (*i) and not Vy, is D ʔmɪlʸ *foot*: Mo hìmé *his foot*, W ʔmɪʔ.)

In a somewhat oblique manner, Diegueños seem to be aware of the intermediate position 'Campo' holds between Diegueño and the Yuman languages farther to the east: one informant, who is very keen on tribal traditions, stated that, when the Southern Pacific railroad gave the Yuma tribe free transportation in exchange for the right of way across the reservation, many Yumas moved to Campo during the summer in order to escape from the heat, and that some of them stayed to settle there. While it seems that some migration has indeed taken place to Southern Californian reservations

[8] I have failed to discover any contrast between this Diegueño x and that corresponding to MaYuCo x, MoYaWH h, as suggested by Kroeber (p. 32). It, therefore, seems appropriate to assume merger of the pre-Diegueño development of a Proto-Yuman 'sibilant' with the earlier state of Diegueño /x/, rather than to set up a separate Diegueño phoneme corresponding to MaYuCo š, MoYaWH s. Since, however, in other Diegueño words, *ceteris paribus*, the sound corresponding to the same extra-Diegueño grouping is š, we seem to have to set up, on the strength of the Diegueño evidence alone, two Proto-Yuman 'sibilant' phonemes, the development of only one of which merged with Diegueño /x/ < Proto-Yuman /x/.

during recent years (at Viejas, certain people, though linguistically at least in part assimilated, were pointed out to me as having come from Arizona, and their speech was labeled vulgar, so that one would not like to talk to them), this story clearly offers no explanation for the characteristics of 'Campo': if any language farther to the east is more closely related to 'Campo', it is Cocopa and not a language of the central group like Yuma —none of the phonological features so characteristic of the central languages is shared by 'Campo': no e is found here corresponding to i in the north, south and west, no δ corresponding to y, no v where the other languages have w.

4. Whether 'Campo' is a reasonably unified language or whether several dialects have to be recognized, is too early to determine. A much greater number of informants will have to be used (as is the case with the other Yuman languages where this problem arises, as with Yavapai or Maricopa), and a thorough understanding of the higher levels of analysis will have to be reached. One case of idiolectic differentiation within 'Campo' (if this is not even a dialectical one) deserves, however, discussion here because it involves a very interesting problem in the change of linguistic structures.

Transition from allophonic to phomenic status occurs when the conditioning factor for the selection of the allophone is no longer present. The conditioning factor is normally conceived of as being some formal feature in the environment of the phone in question.

In one form of 'Campo', now, the following observations could be made: What may be labeled the phoneme /r/ has, in the Yuman languages, a rather wide range of allophonic variation: in Havasupai and Walapai a retroflex t or d or an occasional flapped r, in Yavapai, flapped, trilled, or fricative r; in the other languages, particularly the last two variants are common. The choice of allophones seems rarely idiolectically conditioned; rather does speed, preciseness, etc., seem to play a part. Nowhere could a contrast, on the microlinguistic level, be observed between any of the variants. In this form of 'Campo', however, the informant insisted on contrasting [wɪrá·w] *sharp* with [wɪžá·w] *hot*, with no interchange of [r] and [ž] admitted. The contrast recurred in [kuɾʔák] *husband*: [kužʔák] *old man*, again with explicit rejection of any interchange. Both in other 'Campo' idiolects and in Diegueño, which happens to have the same lexical items in identical distribution, no contrast occurred. From other Yuman languages it seems quite clear that the usages *old man* and *hot* are the primary ones; the spread toward a wider application did not, however, entail any formal change in Diegueño, in other forms of 'Campo', or—for *old man* > *husband*—in Cocopa. In the one idiolect studied, on the other hand, one allophone has become associated with one aspect of the meaning, the other allophone with the second, and since they are no longer interchangeable, the two 'allophones' are now in contrast and therefore phonemes. It is, then, through an interruption in the continuum of meaning, and not through any changes in the formal environment, that the former allophones have gained their new status. Once the allophonic distribution has given way to phonemic contrast in one instance, all other occurrences are subject to redistribution over the two new phonemes. From the scant information at my disposal, both cases with /r/ and /ž/ can be cited, cf. ksá·rk *left (hand)*: tɪmʔúž *full*; no statement on the relative frequency of the two new phonemes is possible at this time. The opposite result of the redistribution upon the phonological shape of the 'primary' and 'derived' members of the pairs *old man*: *husband* and *hot*: *sharp* is noteworthy; it seems to indicate that the alignment with either of the new phonemes was entirely arbitrary, with meaning content not playing

any part in the selection of phonological features beyond forming the basis for the contrast as such.

5. One reason for going into discussions of detail at this early time has been to demonstrate, in a quite informal manner, how closely related to one another the languages of the Yuman group appear to be. If, as has been done here, in treating a problem chosen almost at random, relevant information can be obtained from virtually every idiom of the group, without much distortion to be eliminated, then it is reasonable to assume that at every step of analysis in the detailed study of a single Yuman language we may profitably look for parallels in another language that might confirm or qualify our results. There is, therefore, justification for carrying on concurrently step-by-step analyses of individual idioms of this group and comparative evaluations of the lower levels in analysis attained for every member of the group. It is to be expected that a satisfactory preliminary statement on comparative Yuman phonology can be made once a satisfactory phonemic analysis of the member languages has been performed, regardless of whether, at that time, a complete statement on morphology can be made about all these languages taken separately. The only demand to be made on a concomitant descriptive and comparative treatment of a group of languages is that comparison be not undertaken on a level of analysis higher than that previously mastered descriptively.

PHONEMES OF OCAINA (HUITOTO)[1]

ARLENE AGNEW AND EVELYN G. PIKE
SUMMER INSTITUTE OF LINGUISTICS

1. Consonants
2. Vowels
3. Tone
4. Distribution of phonemes

1. There are twenty six consonant phonemes in Ocaina which contrast as to stops, fricatives, assibilants and nasals. The nasals, perhaps, exhibit the most interesting sound features of the language in that they contrast as to fortis and lenis in their three points of articulation. The fortis feature may be actualized either as length or as tense articulation or as both. (The fortis phonemes will be symbolized with capital letters.)

/Ñ/ and /ñ/: /pĩisoÑo/ *you are saving*; /pĩisoño/ *you are digging*.

/N/ and /n/: /haʔĩNa/ *he spoke*; /haʔĩna/ *hurry*.

/M/ and /m/: /baxoMa/ *remedy*; /ohǫ́ma/ *a youth*.

The stops contrast as to voicing, and the voiceless stops contrast in five points of articulation: bilabial, alveolar, alveopalatal, velar and glottal. The voiced stops contrast in all but the glottal point of articulation. The stops tend to be phonetically long or fortis word medial. The longer or fortis allophone is frequently preceded by a slight aspiration, which when following nasalized vowels is actually a voiceless nasal of the same point of articulation as the stop. The voiced alveolar stop has a quick, flap type of articulation.

The alveopalatal phonemes of Ocaina are another phonetic feature of interest. There is an alveopalatal phoneme in each of the eight series of consonants. They are particularly noticeable for the high, front vocoid timber which, in part, characterizes them. In some instances this is observable as a high front vocoid off glide particularly in the stops /tʸ/ and /dʸ/. There are no monosyllabic consonant clusters so this off glide cannot be interpreted as a /y/ phoneme. Because this glide carries no contrastive pitch as do all vowel phonemes, it cannot be interpreted as a vowel. The stop-high front vocoid glide sequence is, therefore, treated as a unit phoneme.

/p/ and /b/: /opaahị/ *leaves*; /obaahĩ/ *palm nut*.

/t/ and /r/: /mata/ *laugh*; /mara/ *gummy*.

/tʸ/, /dʸ/, and /t/: /tʸadʸóónĩÑo/ *you are eating meat*; /tʸatʸóónĩño/ *you are turning the meat over*; /tʸapóóta/ *you like it*; /tʸapóótʸa/ *you straighten it*.

/k/, /g/ and /ʔ/: /kĩ́rama/ *macaw*; /gĩ́rĩko/ *fish*; /hogąąko/ *palm wood*; /hogaaʔo/ *lizard*; /mokǫǫñǫ/ *you are adding a spice*; /moʔǫǫñǫ/ *you folks are bearing it*.

The fricatives exhibit the same contrasts as the stops except that the voiced fricatives contrast in only two points of articulation: the bilabial and the alveopalatal. The alveolar and alveopalatal fricatives are grooved. The voiceless velar fricative and the voiced fricatives each have an occasionally occurring lenis affricated allophone and the voiced alveopalatal fricative has, in addition to the affricated allophone, one

[1] Ocaina is spoken by a group of people living on the Yaguas Yuca, Ampiyacu, Algodon, and Putumayo Rivers in north eastern Peru. They tell of being transported to Peru from Colombia by landowners between the years of 1930 and 1935 and estimate that they, in Peru, number about two hundred.

The data in this paper were collected during 1955 and 1956 by Arlene Agnew, supplemented by data collected previously by Ilo Leach and Beverly Holcomb. Ocaina belongs to the Huitoto family of languages, listed by Norman McQuown, Indigenous Languages of Latin America, American Anthropologist, Vol. 57, No. 3 Part 1, June 1955.

which has very little friction and grooving. /x/ is rounded before /o/. /o/, /h/ and all the labial consonants are followed by a slight high back vocoid glide preceding /ï/.

/p/, /ƀ/, /p/, and /b/: /tᵛapïïra/ *you open it*; /tᵛabïïro/ *hill*; /pïrąąñą/ *white*; /pïraaži/ *yucca leaf*; /modᵛabóóhoʔta/ *your two foreheads*; /modᵛabóóhoʔta/ *your two tables.*

/s/, /š/, and /ž/: /xooso/ *mine, as of an avocado*; /xoošo/ *mine, as of a house*; /xoožo/ *mine, as of a mosquito net.*

/x/, /h/, /k/, and /š/: /haxážoʔxa/ *he has cut it*; /hahažobï/ *he is resting*; /hakažoho/ *his parrot*; /koNǫxo/ *I drink*; /koNǫšo/ *my bread fruit.*

There are just four assibilants which contrast as to voicing and alveolar versus alveopalatal points of articulation.

/ȼ/ and /ʒ/: /ȼooroƀi/ *a house bird*; /ʒooroʔpi/ *fountain.*

/č/ and /ǰ/: /čičiʔo/ *cricket*; /ǰiǰičora/ *cinamello tree.*

/ȼ/, /č/, /t/, /s/, and /š/: /xooȼo/ *mine, as of cluster of grape-like fruit*; /xoočo/ *mine, as of a pot*; /ïʔxooto/ *almond*; /xooso/ *mine, as of an avocado*; /xoošo/ *mine, as of a house.*

/ʒ/, /ǰ/, and /ž/: /tᵛaʔaʒiiÑo/ *you are hitting it*; /oraǰiiÑo/ *you are playing*; /kodᵛotaǰi/ *I want to get warm*; /kodᵛotaži/ *I'm going to eat.*

Summary of the consonants (from a Chart which lists from left to right: Bilabial, Alveolar, Alveopalatal, Velar, Glottal):

Voiceless Stops p t tᵛ k ʔ and Voiced Stops b r dᵛ g.

Voiceless Fricatives ƥ s š x h and Voiced Fricatives ƀ and ž (bilabial, alveopalatal).

Voiceless Assibilants ȼ č and Voiced Assibilants ʒ ǰ (alveolar, alveopalatal).

Fortis Nasals M N Ñ and Lenis Nasals m n ñ.

2. There are five contrastive tongue positions in the vowels of Ocaina. The high vowels contrast as to front and back tongue positions and both are pronounced with spread lips. The /i/ phoneme embraces allophones from [i] through [ɪᵛ] positions, the lower ones of which may be slightly glided toward a higher tongue position. The allophone with the lower tongue position has been noted especially after /k/ and /h/. The /ï/ phoneme is fronted next to labial sounds and is lowered when it occurs with a low tone between consonants.

The mid vowels contrast in front and back tongue positions: /e/ and /o/. The back one is a rounded vowel and varies freely with higher rounded allophones. The front vowel, most frequently [ɛ], has a set of freely varying allophones consistently lower than the /e/ in the Spanish loan word "docena" meaning 'dozen.' The low vowel /a/ is very comparable to the sounds of the Spanish /a/ phoneme. The more front allophone occurs following alveopalatal sounds. Vowels preceding /ʔ/ are optionally laryngealized. Examples of the five oral vowels are shown in the final syllable of the following words:

/ooʔpi/ *a kind of bird*; /dyoʔóópï/ *a cooking pot*; /oopeʔ/ *your sister-in-law*; /ooʔpo/ *head*; /ooʔpa/ *ahead.*

/i/ and /ï/: /aaʔti/ *it hurts*; /xatï/ *awake.*

/e/, /o/, /i/, and /ï/: /áábeko/ *small fly*; /ohĩboko/ *glasses*; /béteʔo/ *small banana*; /bitiʔo/ *small jungle fruit*; /ïïro/ *tomorrow*, /eere/ *little parrot*; /horo/ *walk, as an animal*; /horï/ *to sift.*

/a/, /e/, /o/, and /ï/: /ooMe/ *his frog*; /ooMa/ *your brother-in-law*; /xora/ *to take something down*; /xoro/ *to break*; /maro/ *shell*; /mïro/ *to rob.*

The oral vowels contrast with nasalized vowels, except /e/ which has no nasalized counterpart; it is a very infrequently occurring vowel.

/ï/ and /ĩ/; /o/ and /ǫ/: /ïïʔta/ *daughter*; /ĩĩʔta/ *many worms*; /nïmǫhį/ *much fruit*; /nį̃mohį̃/ *refuse.*

/a/ and /ą/: /tᵛaháhiʔ/ *fell it*; /tᵛahąhi/ *it is already ripe.*

/i/ and /į/; /o/ and /ǫ/: /dᵛoʔǫǫhį/ *summer*; /dᵛoʔóóhi/ *hot water.*

3. There is contrastive pitch in Ocaina

as seen in the following words (the acute accent over a vowel indicates high tone, whereas a vowel without such an accent indicates low tone):

/xóótʸo/ *your game bird*; /xootʸo/ *game bird*; /tʸáxotʸo/ *I sell it*; /tʸaxotʸó/ *I need to sell it*; /ñonǫ́ǫha/ *kind of pod*; /ñonǫǫha/ *wood worms*.

The following tone groups illustrate the two contrastive tone levels. In two syllable words there is contrast between high-low and low-low as illustrated by two samples from each group:

/tóto óxï/ *big grandmother*; /ihï ása ǒxï/ *he cooks a lot*; /koño óxï/ *big bird*; /ihï motʸo óxï/ *he sews a lot*.

In three syllable words there is contrast between low-low-low; low-high-low and high-low-low:

/dʸaraga óxï/ *a big cut tree*; /porótʸa óxï/ *big bean*; /kážoho óxï/ *big parrot*.

Low tone is optionally progressively lower in pitch from the beginning to the end of the utterance. This is especially apparent in a sequence of like vowels. Palatalized consonants tend to have a slightly rising pitch glide from the consonant to a following high tone vowel and a fast falling pitch glide to a following low tone vowel. Low tones are much more frequent in occurrence than high tones.

Tone is involved in the morphophonemics of the language: maaxa ~ mááxa- ~ maaxá- *walk*; ihï maaxa óxï *he walks a lot*; ihï mááxayo óxï *he is walking a lot*; ihï maaxána óxï *he walked a lot a long time ago*.

The first and third morpheme alternates vary freely in the following phrase: ihï maaxaʔxo óxï, ihï maaxáʔxo óxï *he can walk*.

Tone also frequently indicates items such as mode, tense, person, etc. as: ihï óxï hóórorï *he must crawl like an animal a lot*; ihï óxï hoorórï *he is going to crawl like an animal a lot*; pïïxo *axe*; pïïxo *your axe*.

4. There are the following types of syllables in Ocaina (for this section on vowel distribution, it is usually the first syllable of the example given that is illustrative of the point in question.): V amǫ́ǫ́xǫ *a large ant*; CV našoora *his little wheel*; VC áʔtʸaba *hen*; CVC noʔxáájo *dreamed*. To form a word, up to as many as nine or ten syllables may combine in any order with the one restriction that no more than two vowels may occur in sequence.

Each of the nine Ocaina vowels occurs in each syllable type, except that /e/ and /į/ do not occur in the VC syllable.

V: /a/ agïïMa *worm*; /o/ orootʸo *nigua*; /ï/ ïbaha *lichen*; /i/ ihąąma *fish*; /e/ eere *small parrot*.

CV: tatááko *small frog*; toxooma *shushupi snake*; tïrooho *marmoset*; títʸooro *a plant fibre*; pežiiʔi *small toucan*.

VC: /a/ aʔtiitʸo *tooth*; /o/ oʔkáápe *rope*; /ï/ ïʔžóóga *foot*; /i/ naƀiiʔta *join*.

CVC: baʔxóóma *remedy*; noʔxáájo *dreamed*; pïʔxóózo *chain*; haróti? *people*; /e/ peʔxooti *fish hook*.

V: /ą/ ąNį̈ko *gnat*; /ǫ/ ǰíǫhį *sand*; /į̈/ į̈ʔta *many worms*; /į/ ʊ́ʸáʔąį *throw it out*.

CV: omą́ą́hačo *barbasco stick*; mǫhǫǫʔï *kind of fish*; tį̈tʸaƀóko *button*; hïMohį *pus*.

VC: /ą/ ą́ʔxahį *potatoes*; /ǫ/ ǫ́ʔǫño *you are swimming*; /į̈/ į̈ʔta *many worms*.

CVC: pąʔko *river frog*; hǫʔxo *dog*; kïʔtʸopi *fermented yucca*; /į/ hïMohį? *pus, object*.

Each consonant phoneme occurs in the CV syllable whereas all but the fortis nasals and /š/ occur initially in the CVC syllable.

/ʔ/ is the only member of a special consonant class in that it is the only consonant that occurs following a vowel of the same syllable.

The only limitation in distribution of specific phonemes as regards the word is that the fortis nasals, alveopalatal fricatives and /r/ do not occur word initially.

The only consonant sequences that occur in Ocaina are /ʔ/ followed by one of the following twelve consonants: t, tʸ, k, b, p, s, š, x, ž, ɸ, č, m, and ñ: ï̈ʔto *grandfather*; xóóʔtʸo *deer*; į̈ʔko *wood worm*; ǫǫóʔba *how many*; oʔpóxora *hat*; tʸahóóʔsaño? *are you buying*; bááʔšotʸa *old basket*; oʔxóótʸo

star; ï?žóóga *foot*; ha?ɸópiñǫ *he is painting*; a?čoko *mosquito*; pó?tʸa?mï *straight ahead*; ąą?ño *bittern*.

There seems to be little significant limitation in distribution of specific consonants preceding any of the vowels. Each of the consonants precedes /a/ and /o/. Most of the consonants precede /ǫ/. The alveolar and alveopalatal fricatives (except /s/) and affricates do not precede the vowels /ï/ and /į̈/ but any other lack of possible occurrence seems to be entirely sporadic or in proportion to the frequency of occurrence of each phoneme rather than a structural limitation.

The vowel sequences found to date are of two vowels. The two vowels are either both oral or both nasalized. Among the more frequently occurring vowel sequences are those of like vowels. These phonetically long vocoids are considered to be a sequence of vowels rather than one long vowel phoneme because these phonetically long vocoids parallel in distribution sequences of two unlike vowels which may be seen in the following list of all the vowel sequences that have been found in the language.

The phoneme /o/ has the widest distribution in vowel sequences in that it precedes four of the five oral vowels, including itself: hoá?xa *work*; poï *mouth*; sóeño? *is it breaking*?; tʸa hanooko *he is touching it*.

The phoneme /a/ precedes /i/ and /ï/ as well as itself: háidʸa *salt*; hïbáïxo *butterfly*; haátiɸo *mink*.

/i/ precedes /o/ in addition to itself: bió?xa *fariña*; ïno?xaɸíïbo *cot*.

/e/ precedes /i/ and itself: heíbi *cocoa*; eere *small parrot*.

/ï/ precedes only itself: ïïro *tomorrow*.

The nasalized vowel sequences are fewer in number and less frequent in occurrence than the oral vowel sequences. /ǫ/ has the widest distribution in that it occurs preceding /i/, /ą/, /į̈/, and /ǫ/: hǫ̨ixo *monkey*; móąño? *is he twisting*?; nǫ̨ï *bat*; anïtʸǫǫ́ko *snake*.

/ą/ precedes /į/, /ǫ/, and /ą/: tʸa?ą́įño? *is he throwing it out*?; tʸa?ą́ǫño? *is he taking it out*?; hogąąko *palm wood*.

/ï/ precedes /į̈/: ïį̈?ta *many worms*.

/į/ precedes /į/: kahíįko *squirrel*.

Further evidence that phonetically long vocoids are phonemically a sequence of two vowels is their tone characteristics. The long vocoids may have not only level pitch, but also rising or falling pitch glides which parallel the high-low and low-high tone sequence of two unlike vowels.

$\acute{V}_1\acute{V}_1$: háágoodáži *they want to talk*; ïï?to *grandfather*.

$\acute{V}_1V_1$: ihï áaži *he wants to climb*; hóo?tʸo *sell*.

$V_1\acute{V}_1$: haátiɸo *opossum*; anïtʸǫǫ́ko *snake*.

V_1V_1: xootʸo *game bird*; tʸátóbïïra? *pour out*.

$\acute{V}_1\acute{V}_2$: hïbáïxo *butterfly*.

$\acute{V}_1V_2$: ihï áïño *he is climbing*; oï *brother-in-law*; háidʸa *salt*.

$V_1\acute{V}_2$: hoá?xa *work*; heíbi *coca*.

V_1V_2: tïïroï *bird*; hǫ̨ixo *monkey*.

THE FREQUENCY OF CONSONANT CLUSTERS IN CHONTAL

KATHRYN C. KELLER AND SOL SAPORTA
INDIANA UNIVERSITY

1. The problem
2. The data in Chontal
3. Problems and assumptions
3.1. Distinctive features and the differences between phonemes
3.2. Clusters with liquids and laryngeals
3.3. Chance
4. Conclusions

1. In a recent article[1] it was suggested that a statistical analysis of the relative frequency of consonant clusters would reveal an 'other-than-chance' distribution. Specifically, it was hypothesized 'that the average frequency of a consonant cluster is a function of the difference between the phonemes in the cluster: low frequencies are expected for clusters which are either extremely similar or extremely dissimilar; high frequencies are expected for clusters which are at neither extreme' (25). The results based on texts in Spanish and English seemed to substantiate the hypothesis: 'the distribution of the average frequencies of clusters [when frequency is plotted against similarity] tends in both languages to follow a normal curve' (29).

The present study attempts to test the hypothesis further on the basis of Chontal, a Mayan language spoken in Tabasco, Mexico;[2] and, in addition, on the basis of the Chontal data, re-examine some of the assumptions involved, and discuss some of the problems.

2. The corpus used consisted of texts comprising a total of approximately 72,000 phonemes, recorded from two different informants. Some were story texts and some were descriptions of native customs. Clusters across (morphologically determined) word boundaries were not included.[3]

The procedure in this study parallels that used with the English and Spanish materials. An analysis into binary distinctive features[4] was made of the Chontal phonemes. The distinctive feature analysis for those consonants which are being considered in this study appears in Table 1. The difference between consonants was then calculated on the basis of their distinctive features. A direct contrast (represented in the table by + versus −) is considered two units of difference and a 'semi-contrast' (+ or − versus non-pertinent) is considered one unit of difference. The analysis is restricted to those sequences which are consonantal and non-vocalic. Combinations including (a) vowels (i.e., phonemes which are − consonantal and + vocalic), (b) liquids (i.e., phonemes which are + consonantal and + vocalic) and (c) laryngeals (i.e., phonemes which are − consonantal and − vocalic) are not included (see discussion below in

[1] Sol Saporta, Frequency of Consonant Clusters, Lg. 31.25–30 (1955). We have profited from discussions with H. V. Velten and George Herzog.

[2] For a previous analysis of Chontal data, see Kathryn C. Keller, The Chontal (Mayan) Numeral System, IJAL 21.258–275 (1955).

[3] Possibly the unit chosen as a basis for the count is a factor in the relative frequency of consonant clusters. It would be interesting, for example, within any given language, to determine frequencies within (1) syllables (2) morphemes (3) words (4) phrases (5) complete utterances. One might speculate that the higher the order of the unit, the more random (i.e., predictable by chance) the frequencies of the clusters. Certainly, the initial consonant of one utterance is independent of the final consonant of the previous utterance.

[4] For a presentation of the theory of binary distinctive features, see Jakobson, Fant and Halle, Preliminaries to Speech Analysis: the Distinctive Features and Their Correlates (Cambridge, Mass., 1952).

TABLE 1—The Distinctive Features of the (non-liquid and non-laryngeal) Consonants in Chontal. All these phonemes are + consonantal and − vocalic

	m	p'	p	b	w	n	c'	c	t'	t	d	č'	č	k'	k	g	s	š	y
grave/acute	+	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−
nasal/oral	+	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−
interrupted/continuant		+	+	+	−		+	+	+	+	+	+	+	+	+	+	−	−	−
strident/mellow							+	+	−	−	−	+	+	−	−	−	+	+	−
compact/diffuse							−	−	−	−	−	+	+	+	+	+	−	+	
strong/weak		+	+	−					+	+	−			+	+	−			
checked/unchecked		+	−				+	−	+	−		+	−	+	−				

Table 2—Chart Representing Pertinent Distinctions

	grave				acute						
					diffuse			compact			
nasal	m				n						
continuant	w				y						mellow
					s			š			strident
oral											
interrupted	p'	p	b		t'	t	d	k'	k	g	mellow
					c'	c		č'	č		strident
	check.	unche.			check.	unche.		check.	unche.		
	strong		weak		strong		weak	strong		weak	

TABLE 3—Units of Difference

	m	p'	p	b	w	n	c'	c	t'	t	d	č'	č	k'	k	g	s	š	y
m	0																		
p'	5	0																	
p	5	2	0																
b	4	3	3	0															
w	3	4	4	3	0														
n	2	7	7	6	5	0													
c'	8	5	7	6	7	6	0												
c	8	7	5	6	7	6	2	0											
t'	9	4	6	7	8	7	3	5	0										
t	9	6	4	7	8	7	5	3	2	0									
d	8	7	7	4	7	6	4	4	3	3	0								
č'	8	5	7	6	7	6	2	4	5	7	6	0							
č	8	7	5	6	7	6	4	3	7	5	6	2	0						
k'	9	4	6	7	8	7	5	7	2	4	5	3	5	0					
k	9	6	4	7	8	7	7	5	4	2	5	5	3	2	0				
g	8	7	7	4	7	6	6	6	5	5	2	4	4	3	3	0			
s	7	8	8	7	4	5	3	3	6	6	5	5	5	8	8	7	0		
š	7	8	8	7	4	5	5	5	8	8	7	3	3	6	6	5	2	0	
y	6	7	7	6	3	4	6	6	5	5	4	6	6	5	5	4	3	3	0

3.2). A graphic representation of the pertinent distinctions is shown in Table 2. Table 3 shows the difference, in units, between each pair of consonants.

The occurrent clusters and their frequencies are indicated in Table 4. Numbers in column A denote the difference, in units, between members of a cluster. Column B

TABLE 4—The Clusters and Their Frequencies

A	B	C	D	E	F
0	bb 1	19	1	1	.05
1		—	—	—	.—
2	kt k'k tk nm 57 6 5 2	22	4	70	3.18
3	šď' čk šď 65 53 14	40	3	132	3.30
4	mb k't kp ny yn pt t'k pk p'k' 66 54 10 9 4 4 4 1 1	44	9	153	3.48
5	mp' nš šn mp šc yk čt nw ns yt ky kď' yk' šc' č'k k'č 109 25 23 16 13 10 4 3 3 3 2 2 1 1 1 1	62	16	217	3.50
6	nd šk nď' nc' nč st nb čb c'b pk' ng ym yb p'k bč tp' č'b 264 107 17 14 9 9 7 6 5 5 2 2 2 2 1 1 1	58	17	454	7.83
7	nt nk' ms nk k'b tb kb kn šb nt' šm sb bt bs sm t'b c'k yp' č't k'n bk 394 116 89 79 37 28 18 17 16 13 12 10 9 6 5 3 3 2 2 1 1	72	21	861	11.96
8	sk k's št ps šp šp' ks kw 135 50 36 14 7 4 4 1	36	8	251	6.97
9	k'm mk mt km mt' 12 8 7 6 1	8	5	34	4.25
		361	84	2173	

lists the clusters in order of frequency. Column C gives the number of theoretically possible clusters for each degree of difference, (19 × 19 =) 361 in all. Column D gives the number of different clusters of each kind which actually occur. Column E indicates the total frequency of these clusters, and column F gives the average number of clusters in each category (column E divided by column C).

The distribution of the average frequencies (column F), although skewed, has only one peak, near the mean.

3.1. It is obvious that the results tabulated above (Table 4) are only as valid as the distinctive feature analysis on which they are based, since it is this analysis which determines the difference between consonants. The validity of the distinctive feature analysis must be investigated in two directions. First, does the analysis as presented accurately reflect the Chontal data? Second, is a distinctive feature analysis the most valid way of determining the difference between consonants?

Regarding the first question, in the absence of clear-cut spectrographic data, the distinctive feature analysis above represents an attempt to convert the articulatory data to the system presented in the Preliminaries. Where the conversion is not automatic, other criteria, in part arbitrary perhaps, have been used. The analysis of /h/ may serve as an illustration. The phonetic range of this phoneme is from a glottal spirant [h] to a weak velar spirant [x], [x̯]. It may be considered a consonant similar to /w/ and /y/ (which have fricative allophones), or a laryngeal, similar to /ʔ/. Its distribution parallels that of /ʔ/ and it is therefore analyzed as a laryngeal (i.e., − consonantal and − vocalic).[5]

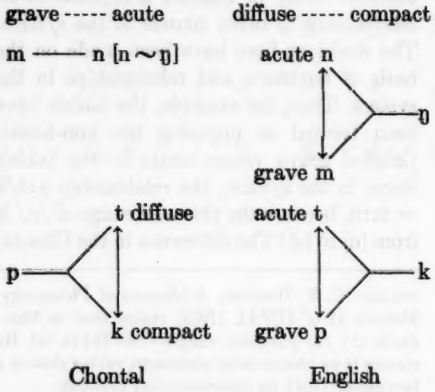
The question arises whether there is any clear-cut way of deciding whether a particular sound is consonantal or vocalic on an absolute basis,[6] or whether it depends on its relationship to other sounds in the system. The decisions here have been made on the basis of contrasts and relationships in the system. Thus, for example, the labials have been treated as opposing the non-labials (labeled grave versus acute in the table), since, in the system, the relationship p:t/k = m:n, because the phonetic range of /n/ is from [n] to [ŋ].[7] The difference in the Chontal

[5] Douglas Taylor, Phonemes of the Hopkins (British Honduras) Dialect of Island Carib, IJAL 21.233–41 (1955) reports a phoneme /h/ which ranges from velar spirant to weak aspiration. He analyzes it as consonantal and non-vocalic. C. F. Hockett, A Manual of Phonology, Memoir 11 of IUPAL (1955) states that in Mandarin the /x/ phoneme ranges from [h] to [x]. He classes it as 'dorso-velar obstruent rather than as a laryngeal' (107) on distributional grounds.

[6] See Hockett, Manual, p. 38, for a discussion of the borderline between vocoids and contoids.

[7] It is not clear how a distinctive feature analysis enables the analyst to discover the phonemes of a language (but see Morris Halle, The Strategy of Phonemics, Word 10.209 (1954) who maintains that 'the analysis . . . permits us to establish the inventory of the phonemes of the language'), nor is it clear whether a particular feature must be equally distinctive to all the allophones of a phoneme. Halle suggests that 'the method of minimally different sets avoids the difficulties connected with allophones . . . The much discussed problem of whether we should consider the English [h] and [ŋ] as one or two phonemes does not arise at all, because the answers from analyzing *hill* in a minimally different set like *hill, pill, till, bill*, etc. differ completely from those obtained by analyzing *sing*, in a set like *sing, sin, sick, sit*, etc.' Such a method would cause difficulties, it seems, in the establishing of phonemes in a language (like Spanish) where almost all phonemes may occur in initial position, but only very few in final position. The answers from analyzing *dan* in a set like *dan, pan, van, tan*, etc. would differ from those obtained by analyzing *ved* in a set like *ved, ver, ves, ven*. (The only consonants occurring in final position are /d r s n l x/.) The initial [d] in *dan* could not then be classified with the final [d] of *ved*. Thus, in the Chontal example, [m] contrasts with [n] initially and intervocalically, but with [ŋ] in final position; nevertheless, /m/ has been analyzed as distinctively grave and /n/ [n ~ ŋ] has been analyzed as distinctively acute. In a language like English, however, where /m/, /n/ and /ŋ/ are different

system (as set up here) and the English system (as set up in the Preliminaries) can be represented as follows:



Regarding the problem of the validity of any measure of difference between phonemes which is based on a distinctive feature analysis, it is apparent that the final solution determines the number of units of difference. Consider a hypothetical problem where one is concerned with the relative differences between the members of a set of labials /p b m/. There are at least the following theoretically possible analyses:

(1)	p	b	m
Voiced/Voiceless	−	+	+
Nasal/Oral		−	+

The differences, according to the above, are p/b = 3; p/m = 3, b/m = 2; b is closer to m than it is to p.

(2)	p	b	m
Nasal/Oral	−	−	+
Voiced/Voiceless	−	+	

The differences are p/b = 2; p/m = 3; b/m = 3; b is closer to p than it is to m.

(3) (Includes redundant features; in effect, in this example, the analysis would measure the difference between phones rather than between phonemes.)

	p	b	m
Nasal/Oral	−	−	+
Voiced/Voiceless	−	+	+

The differences are p/b = 2; b/m = 2; p/m = 4; b is equally distant from p and m. Obviously, then, the order of presenting the features as well as the inclusion or exclusion of redundant features alters the difference between consonants.

It should be pointed out that in the usual type of distinctive feature analysis an empty space (or zero) in the table (i.e., neither plus nor minus) may be used in different types of situations. In some cases, the empty space could be replaced by a sign which, however, would be redundant. Thus, in (2) above, /m/ is voiced but not distinctively voiced. If in the hypothetical case (2), /m/ had, say, both voiced and voiceless allophones, the empty space would indicate that one allophone was plus, and one minus. In addition it has been suggested that 'a zero appears in a place where the substitution of plus or minus would imply an impossible articulation.'[8] In all cases, however, the zero indicates that the feature is not necessary for the identification of the particular phoneme.

Regarding the measurement of differences, it may be of interest to indicate similarities in another dimension, namely similarities in distribution. Presumably, this could be done in a manner somewhat parallel to that used in obtaining differences by distinctive (phonetic) features. For the particular language, a list of distinctive distributional features could be determined and each consonant checked against it. In Chontal and Spanish, for example, intervocalic position would not be a distinctive distributional feature of consonants, since every consonant may occur in that position; however, final position would be distinctive, since certain consonants do not occur finally. Within the

phonemes, /m/ is distinctively diffuse in contrast to /ŋ/, but distinctively grave in contrast to /n/. See Jakobson, Preliminaries, pps. 40, 43.

[8] Cherry, Halle and Jakobson, Toward the Logical Description of Languages in Their Phonemic Aspect, Lg. 29.39 (1953).

TABLE 5—Clusters in Chontal, Independent of Frequency

	m	p'	p	b	w	n	c'	c	t'	t	d	č'	č	k'	k	g	s	š	y	
m		+	+	+					+	+					+		+			7
p'														+	+					2
p										+				+	+		+			4
b				+						+			+		+		+			5
w																				0
n	+			+	+		+		+	+	+	+	+	+	+	+	+	+	+	15
c'				+											+					2
c																				0
t'				+											+					2
t		+		+											+					3
d																				0
č'				+						+					+					3
č				+						+					+					3
k'	+			+		+				+			+		+		+			7
k	+		+	+	·:·	+				+		+					+		+	9
g																				0
s	+			+						+					+					4
š	+	+	+	+		+	+	+		+		+	+		+					11
y	+	+		+		+				+				+	+					7
	6	4	3	13	2	4	2	1	2	11	1	3	4	4	14	1	6	1	2	84

same language, two consonants which are very similar according to distinctive phonetic features, may be very different according to distributional features. The nasals /n/ and /ŋ/ in English are very similar as determined by their distinctive phonetic features, but their distributions are quite different.[9] The same is true for the pair /š/, /ž/. The extent to which classifications based on distributions correlate with classifications based on articulatory and/or acoustic data will probably vary from language to language.[10]

Table 5 indicates the nature of the 84 clusters which occur in the Chontal texts examined. The first member of the cluster is indicated by the phoneme listed at the left of each row; the second member of the cluster is indicated by the phoneme listed at the top of each column. The data indicate, for example, that the most frequent phonemes occurring as second members are the strong (non-labial) stops, /k/ (14), /t/ (11) and the weak (labial) stop /b/ (13); /d/ and /g/ occur only after /n/; /p/ after /m/ and after /k/ and /š/.[11]

3.2. As indicated above, clusters involving /r l h ʔ/ have been excluded in this analysis. These clusters are quite common in Chontal, the total reaching 2353 clusters in the ma-

[9] Sapir, Sound Patterns in Language, Lg. 1.37–51 (1925) discusses relationships of consonants based on other than 'organic and acoustic grounds.' Regarding the /ŋ/ of English, in its relationship to the other nasals, he maintains that 'in spite of what phoneticians tell us about this sound (b:m as d:n as g:ŋ), no naive English-speaking person can be made to feel in his bones that it belongs to a single series with m and n. Psychologically it cannot be grouped with them, because, unlike them, it is not a freely movable consonant' (49).

[10] For pertinent discussions on the classification of phonemes on the basis of distribution, see Bloomfield, Language (1933) pps. 130–134; and, recently, Kenneth L. Pike, Language in Relation to a Unified Theory of the Structure of Human Behavior, Part II (Glendale, Calif., 1955) pps. 25–30; Eli Fischer-Jørgensen, On the Definition of Phoneme Categories on a Distributional Basis, Acta Linguistica 7.8–39 (1952); Hans Vogt, Phoneme Classes and Phoneme Classification, Word 10.28–34 (1954).

[11] It is worth noting that the distributional data seems, in this case, to corroborate the division of the stops into labials (grave) and non-labials (acute).

terial examined. The distinctive feature analysis indicates the marginal status of these phonemes. Since the two features of vocalic/non-vocalic and consonantal/non-consonantal are primary, it is almost inevitable that the four resulting classes of phonemes (vowels, liquids, laryngeals, and what may be termed full consonants) will be comprised of members whereby almost any full consonant will be maximally different from any liquid and laryngeal. For example, if there are, as in Chontal, two laryngeals and nineteen full consonants, it is clear that there can be only one further feature which will be distinctive to laryngeals, but at least five[12] additional features required for the full consonants. The result, in general, will be that members of this class will be further apart from laryngeals than they are from one another. Given this fact about differences between phonemes as measured by distinctive features, it seems likely that the inclusion of sequences involving laryngeals or liquids (where such sequences are common) will almost always skew the curve, and that the particular hypothesis being tested here could not be valid. This is not to imply that there is no 'lawfulness' in sequences involving these phonemes, but only that they need not conform to general statements which seem to apply to other consonants.[13]

3.3. The question arises whether the results might not have been expected by chance. The argument may perhaps be best illustrated by an analogy. Imagine eight coins being tossed so as to yield a sequence of heads and tails, say, H H T H T T T H. If we now toss the same eight coins again and compare the second sequence with the first we would not expect exactly the same sequence, H H T H, etc. (minimal difference), nor would we expect the exact opposite, T T H T, etc. (maximal difference); what we would expect, purely by chance, is a sequence in which some coins were the same and some were different. Given a large number of comparisons of adjacent sequences, we would predict a normal distribution, i.e., low frequencies of sequences with no difference or with a very large difference, and high frequencies of sequences at neither extreme. For example, if we represent the difference between the two sequences by a number from 0 (exactly the same) to 8 (exactly the opposite), then, given 256 such comparisons, we would expect, purely by chance, the following fre-

[12] The MINIMUM number of binary features necessary to identify n phonemes is $\log_2 n$. The $\log_2 19 = 4.+$; five is the next higher integer. The MAXIMUM number of distinctive features, i.e., features which serve to distinguish at least one phoneme from the rest, is n − 1. The formula n(n − 1)/2, cited by Halle (see fn. 7) indicates the number of 'differences' between n phonemes. (See W. F. Twaddell, On Defining the Phoneme, Language Monographs 16, Baltimore, 1935, p. 53, who refers to the 'maximum number of significant phonological differentiations.') In Halle's example, given the phonemes /p/, /t/, /f/, /s/, /b/, /d/, /v/, /z/, we would need only seven statements to IDENTIFY each one. Answering the 28 'questions of the form "Is the phoneme under consideration /p/ or /t/, /p/ or /b/, etc.?"' (203) would yield the maximum number of differentiations between eight phonemes. In order to identify the eight phonemes, however, we would need merely to answer seven questions of the form 'Is the phoneme under consideration /p/?' In short, given any three elements A B and C, two of them must share one feature: B and C share the feature 'non A-ness.' Thus, given eight phonemes, there are at most seven distinctive features, but, at most, 28 differentiations (or minimal pairs).

[13] A similar conclusion regarding laryngeals is implicit in the statement by Eunice V. Pike, Phonetic Rank and Subordination in Consonant Patterning and Historical Change, Miscellanea Phonetica, II.25–41 (1954) that the 'low phonetic ranking of /h/ and /ʔ/ may be the reason that, as a pair of consonants, they so frequently pattern differently from other consonants, and at times are so closely allied to the structure of the vocalic nucleus, such that, for example, VʔV may be similar in distribution . . . to a single vowel.' (30) See, too, Hockett's analysis of /r h ʔ/ in various languages as part of a complex peak (i.e., vocalic nucleus) rather than part of the coda or interlude (i.e., final or medial consonant cluster), in the Manual pps. 72–80.

quency distribution:

Difference between two sequences	Number of occurrences[14]
0	1
1	8
2	28
3	56
4	70
5	56
6	28
7	8
8	1
	256

The problem, then, is whether the binary distinctive features would behave like coins, so that the difference between any two consonants in a cluster (i.e., two adjacent bundles of pluses and minuses) would follow the same distribution as the difference between two sequences of coins (i.e., two adjacent bundles of heads and tails). If the answer is yes, then the results are predictable by chance.

There seems to be an important difference between the tossing of the coins and the arrangement of the distinctive features. For a normal distribution, the coins must be unbiased, that is, for each coin the probability of H is equal to the probability of T ($= \frac{1}{2}$). This is not the case with distinctive features, where, for example, in the inventory, there are many more oral consonants than nasal.[15] This factor, the lack of equiprobability (of + or −) within any one feature would tend to disrupt the normal distribution. It seems reasonable to assume, then, that the results indicated above (Table 4) would not be expected by chance alone.[16]

4. The results, based on a corpus of 72,000 phonemes, tend, in general, to support the hypothesis that the frequency of a consonant cluster is a function of the difference between the members of the cluster. These results, furthermore, are not predictable by chance alone. However, the hypothesis excludes consonants which are either liquids or laryngeals. There remains, too, some question as to the most valid technique for determining the differences between phonemes. As currently developed, the distinctive feature analysis provides a rough measure of these differences.

---

[14] The values in this column are the coefficients of the terms which result from the expansion of the binomial $(H + T)^8$; e.g., $1H^8 + 8H^7T + 28H^6T^2 + \ldots + 1T^8$. The number 256 is the sum of these coefficients. For a description of the binomial distribution, see, e.g., Yule and Kendall, Introduction to the Theory of Statistics[11] (London, 1937).

[15] Over a text, the plus probability of a feature is the sum of the probabilities of phonemes showing a plus in the inventory. Thus, the probability, in Chontal, of + nasal is the probability of /m/ plus the probability of /n/. Cherry, Halle and Jakobson (see fn. 8) give the probabilities of plus or minus for several distinctive features in Russian, over a corpus of approximately 10,000 phonemes.

[16] In addition, within any one sequence, the coins are independent, i.e., the occurrence of a head on the first coin does not affect the probability of getting a head on the second or third coin; this is not the case with distinctive features. The most obvious case is the non-independence of the features vocalic/non-vocalic and consonantal/non-consonantal. Thus, given the fact that a particular phoneme is, say, non-vocalic, it is likely (with greater-than-chance probability) that it is also consonantal. In other words, there are significantly more full consonants (− vocalic + consonantal) than laryngeals (− vocalic − consonantal). However, these features would not affect our data.

A STATISTICAL STUDY OF CONSONANTS IN NEW WORLD LANGUAGES
I: INTRODUCTION

JOE E. PIERCE
ANKARA, TURKEY

0. Introduction
1. The sample
2.1. Orthography
2.2. Procedure

0. In a recent paper C. F. Voegelin attempted to type phonological systems on the basis of the number of series of stop consonants.[1] The purpose of this paper is to test statistically the validity of such a grouping. Many points of divergence could be selected on which to classify languages, and the number of series of stops is only one of these. The question then arises does this grouping correlate with other features of the phonological system, thus giving the group meaning, or are they simply a convenient way of classifying consonant systems?

The results of the investigation bring out some very interesting things about the languages of the New World. However, almost as important as the results obtained is the method employed. Until recently the application of sampling was not feasible in linguistics. However, the field has expanded now to the point where there are many field workers turning out descriptions of new languages all the time. This means that in the very near future there will be sufficient good descriptions of languages throughout the world to make sampling meaningful.

The purpose of this investigation is simply to examine a sample of languages, representing a wide range of structural types and geographic locations in the New World, and to see what can be learned about languages from this type of study. The approach is strictly descriptive.

1.1. The sample consisted of all American Indian languages described in the following publications, prior to June 1954:

1 American Anthropologist
2 American Anthropological Association Memoirs
3 Annual Reports of the Bureau of American Ethnology
4 Anthropos
5 Bulletins of the Bureau of American Ethnology
6 Bulletins of the Hervas Laboratories of American Linguistics
7 Handbook of American Indian Languages, pt. 3
8 International Journal of American Linguistics
9 Journal de la Société des Americanists de Paris
10 Language
11 Peabody Museum Papers
12 Publications of the American Ethnological Society
13 Publications of the International Congress of Americanists
14 University of California Publications in American Archeology and Ethnology
15 Viking Fund Publications in Anthropology

The New World was selected both because it was the author's field of specialization and because it contained a wide variety of linguistic structures. From the above list of publications, descriptions of languages repre-

[1] C. F. Voegelin, Inductively arrived-at models for cross-genetic comparisons of American Indian languages. Papers from the symposium on American Indian linguistics held at Berkeley July 7, 1951, (University of California Publications in Linguistics 10.1-68, 1954, 27-45).

senting the following linguistic families north of Mexico were extracted: Eskimo, Wakashan, Kutenai, Quileute, Salish, Yurok-Wiyot, Algonquian, Tlingit, Haida, Athabascan, Uto-Aztecan, Tanoan, Kiowa, Zuni, Tsimshian, Chinook, Alsea, Siuslaw, Maidu, Miwok, Yokuts, Costanoan, Karok, Shasta-Achumavi, Yana, Washo, Yuki, Pomo, Salinan, Chumash, Yuman, Keresan, Caddoan, Tonkawa, Karankawa, Atakapa, Chitmaka, Tunica, Nachez-Avoyelles-Taensa, Muskogean, Iroquoian, and Yuchi.

In addition to languages representing the linguistic families listed above the sample contained sixty-five languages from the remainder of the New World. These are: Alakaluf, Amahuca, Amuesha, Aymara, Campa, Carib, Cayapa, Uro-Chipaya, Chatino, Chontal, Coahuilteco, Colorado, Comercrudo, Copainala, Cora, Coroado, Cotoname, Cuicateco, Cuna, Goajiro, Huasteco, Huicol, Ipurina, Kaiapo, Kaingang, Kaskiha, Kota, Kekchi, Kechua, Kraho, Kuduveo, Kuruaya, Maku, Maritino, Matako, Tzeltal, Mazateco, Tojolabal, Mixe, Motilone, Nahuatl, Nambikwara-Tarunde, Ona, Mazahua, Paez, Payas, Sierra Populuca, Puri, Seta, Shipaio, Terena, Tepecano, Tlappanecan, Tokano, Totonaco, Tsotsil, Tule, Waiwai, Witoto, Yunka, Yaghan, Yagua, Zapotec, and Zoque.

The purpose of this investigation was to correlate the various elements of the phonological systems. Since the phoneme is the smallest phonological segment,[2] the first step in the procedure was to correlate phonemes, or more accurately the symbols which represent these phonemic distinctions, with the various types. However, since phonemic descriptions are available for only a small number of languages, (because phonemics is a relatively late development in the science of linguistics) all phonological descriptions from the New World which were readily available (i.e. those found in the publications listed above) were collected; this list is of course not exhaustive.

1.2. The original data were gathered by many different field-workers over a period of approximately seventy-one years.[3] As a result of this, there is a great disparity, not only in the symbols used, but in the phonetic quality represented by the same symbol, when used by different authors. This presents quite a problem, because each symbol had to be examined and a decision made as to just what the phonetic character represented by that symbol was. Then the symbol had to be replaced by one which would represent the same phonetic quality in every description.

To begin the discussion it is proper to consider the quality of the consonant systems used. Since, as stated above, the basic element of the phonological system is the phoneme, it would be ideal to have modern phonemic descriptions of a large number of languages. However, this is not the case, so a method of rapidly phonemicizing was devised. The procedure is described below, but it must not be assumed that this procedure resulted in an accurate phonemic description of every language. These systems are only close approximations of the phonemes of any given language.

The first step in the procedure was to compare all of the descriptions of a given language where more than one were available. The problem in dealing with nonphonemic systems is to determine which sounds are phonemic and which are not. This is not as difficult for the consonants as it might at first seem, because most of the descriptions had statements such as —[m] occurs only by assimilation.[4] This means then that [m]

[2] For a discussion of the problem of segmentation see, Zellig Harris, Methods in Structural Linguistics, pp. 25–41 (1951).

[3] From a description of the Central Eskimo, BAE. 6.409–666 (1884) to Examples of Tone in Crow, IJAL.20.34–6 (1954).

[4] Pliny Earle Goddard, Wailaki Texts, IJAL. 2.78–135 (1923).

occurs only before bilabial consonants, and in some other phonetic form before vowels and non-bilabial consonants. Statements such as the above made possible the elimination of certain symbols from the systems, because they were positional variants. After some symbols had been eliminated on the basis of distribution, the remainder of the symbols, which all field workers had listed, were accepted as phonemes. Those which were listed by only part of the reporters were accepted or rejected on the basis of the author's judgment, which was influenced by his knowledge of the persons who recorded the languages as well as the publication date.

Forty-five languages (or slightly less than 25 percent of the sample) were represented by a single description, which was not phonemic. The quality of these descriptions varied from a single list of symbols to a rather extensive phonetic description of the sounds in the language. These descriptions were used as given, except where distributional statements permitted the combining of some symbols into a single phoneme. Such statements were very rare. It is difficult to say just how near phonemic the descriptions were, but probably in no language would a phonemicization result in the change of more than perhaps four consonants. This is of course an estimate. However, it will serve to indicate that phonemicization of these languages would not appreciably alter the final results. For the entire sample the average number of phonemes per language is approximately nineteen. If five of the consonants in each language represented by a nonphonemic description (i.e. 25 percent of the sample) were changed by phonemicization, which is extremely unlikely, it would mean an alteration of only six percent of the consonants in the languages of the sample. In all probability the change would be of much smaller magnitude even than this.

Of the 176 languages in the sample 56, i.e. 32 percent are listed as phonemic, or the author of at least one description for each language of this group referred to the symbols as phonemes. It would probably be more accurate to label the systems which were derived by the above procedure as near-phonemic, but in this paper the symbols are treated as if they were phonemes. It would be better of course, i.e. the materials would be more consistent, to work only with modern phonemic descriptions. However this would have reduced the size of sample too much, so the above described compromise was worked out.

In order to better judge the accuracy of the derived systems, consider the raw materials again. Of the 176 languages, 45 (i.e. 25 percent) have a single nonphonemic description. This means that 75 percent of the languages in the sample had two or more non-phonemic descriptions by different authors, or there was at least one phonemic description, and in many cases both a phonemic and a non-phonemic description. Thus in a great majority of cases it was relatively simple to assign to any given language a system of consonants which more than one field worker had listed as phonemes. Because of this fact, it is highly unlikely that the conclusions drawn from this investigation would be altered if phonemic descriptions of the languages in the sample were substituted for those used.

2.1. Much has been said to this point about the difficulties involved in reducing the data to a common set of symbols, but nothing has been said about the symbols themselves. Chart I contains the set of symbols used for the phonemic entities in this investigation.

Each symbol in Chart I indicates three articulatory features; (1) the point of articulation, (2) the mode of articulation, i.e. whether the closure is complete (stop), partial (spirant), complete for the oral passage with the nasal passages left open (nasal), or non-existant (semi-consonant), and (3) whether the vocal cords are vibrating or not. The chart is to be read in the usual manner. The points of articulation are written at the

	Bilabial	Labio-dental	Inter-dental		Alveolar Dental					Lateral	Palatal		Velar	Back Velar	Glottal	
Stops																
Voiced	b		ð	dð	d̪	d	dz	ɖ	dʒ	dl	g̪	g		g̱		
Voiceless	p	pf	θ	tθ	t̪	t	ts	ʈ	tʂ	tl	k̪	k	kx	ḵ	kx	ʔ
Spirants																
Voiced	b̥	v				r	z	ɾ	ʒ	l	l̥	ɣ̪	R	g	ɣ̱	
Voiceless	p̥	f					s		ʂ	ɬ		x̪		x	x̱	
Nasals																
Voiced	m	m			n			n						ŋ	ŋ	
Voiceless (Whispered)	M				N									Ŋ		
Semi-Consonants																
Voiced	w				y											h
Voiceless (Whispered)	W				Y											H

CHART I

These are the symbols which were used in the investigation to represent 1) point of articulation, 2) mode of articulation, and 3) the presence or absence of voicing for each phoneme. Other vocalic qualities are indicated by the presence of certain diacritics.

top of the chart and the modes of articulation down the left hand margin.

One final item concerning the chart: the blank spaces indicate only that the phonemes represented by the symbols which would fill those spaces were not found in any languages in the sample. For example /bv/ (a voiced, labiodental, affricated-stop) the voiced counter-part of /pf/ does not occur in any language of the sample. Therefore the spot for this symbol is left blank. This means that every symbol on the chart is represented in at least one language in the sample, and that every phoneme that occurs in any language in the sample can be represented either by one of the symbols alone or by one of the symbols plus a diacritic or diacritics; with three notable exceptions, which will receive special treatment later.

As mentioned above, certain diacritics were used in addition to the symbols on the chart. These marks do not change the point, or mode of articulation, or voicing, but add certain articulatory features to those already denoted by the symbol as taken from Chart I.

First and statistically the most common of these diacritics is the right hand single quote (’), which when placed immediately following a consonant symbol, indicates glottalization of the consonant in question.

Second left handed single quote (‘), which when placed immediately following a stop represents aspiration, or when following a non-stop consonant, excessive aspiration.

Third is a raised /ʷ/, which indicates labialization of the consonant.

Fourth a raised /ʸ/, which indicates palatalization of the consonant symbol.

Next are some diacritics which are placed above or below the consonant symbol, rather than immediately following it. The first of these is a pair of reversed commas (‟) placed beneath a consonant, which indicates fortis articulation.

The remainder of the diacritics, which occur above or below the consonants that they modify, have a very limited distribution, i.e. they occur with a very restricted number of symbols. However, these occur with a relatively high frequency in terms of their appearance in languages of the sample. The first of these is the inverted wedge (ˇ), which when placed over the sibilants (s and z) indicates dorsal rather than apical articulation.

The tilde (˜) has a very limited distribution also, occurring only with the series of

/r/'s, indicating trilled as opposed to flap articulation. Since there are dental, alveolar, and palatal /r/ symbols, and since all of these can be either trilled or flapped, it seemed easier to set this up as a diacritic rather than to create a contrasting series; the latter would of course be possible.

The raised (·) indicates length. This doesn't mean simply phonetic, but rather phonemic length. Whether length was phonemic or not presented no problems in dealing with the consonants.

As always, when diacritics are used, some conventions had to be adopted as to the method of combining them. Quite naturally, the symbol taken from Chart I is the place to begin. There are two types of diacritics described above, those occurring directly above or below the symbol, and those immediately following it.

The first group includes (◌̈), (◌̌), and (◌̃), and need no further explanation, since they are mutually exclusive.

However, the second type, which includes (’), (‘), (ʷ), (ʸ), and (·), sometimes occur simultaneously or near-simultaneously modifying the same consonant, and for the sake of consistency the following conventions were adopted. The symbol from Chart I can be followed immediately by any of the five diacritics above, e.g. k’, k‘, kʸ, kʷ, and k·. However, if more than one of them occur in conjunction with the same consonant, the length mark occurs first in the sequence. Following the length mark any of the four remaining diacritics may occur, e.g. k·’, k·‘, k·ʸ, and k’ʷ. If the sequence includes more than two diacritics (or only two if the length mark is not in the sequence) then the right handed (’) or left (‘) single quote marks (which never occur together) come first, followed by (ʸ) or (ʷ), which are also mutually exclusive. In the case of the affricates the right or left handed single quote, (’) or (‘), comes immediately after the stop element and is followed by the spirant element of the symbol, e.g. /t’s/ stands for a glottalized dental affricated-stop, with apical articulation.

One convention of a slightly different type remains to be explained. This is the use of homorganic nasals placed immediately preceding a stop. These complex symbols represent nasalized consonants.

With the symbols in Chart I, plus the diacritic marks discussed above, it is possible to represent all of the consonants[5] which were described by the various field workers for all languages in the sample except one. This single exception is the phoneme /br/, which occurs only in Osage.[6] It is a unique sound, or sound sequence, which gives the acoustic impression of English (br) or (bl) cluster, depending on the dialect of Osage which is being spoken.

2.2. With all the consonants reduced to a standard set of symbols (cf. Chart I) the next step was to count the frequencies of occurrence for each symbol. Table I gives a complete list of the symbols with the frequency of occurrence for each symbol and a computed percentage of the languages in which that consonant occurs, down to a percentage of occurrence of three percent. All symbols occurring in less than three percent of the sample, i.e. in less than five languages, are listed below without either frequency or percentage:

---

w’, g̱, k‘, m’, k̂’, l’, x̣ʷ, N, l̥ʸ, gʷ, k‘ʷ, mb, nd, ndž k˙, t̥’, ŋ’, k̥‘, dʸ, r̥, W, Y, f’, l̥ʸ, n’ʸ, k̥, t̥, p̥, M, ŋʷ, nx, g‘, hʸ, b, t̥š, k̥ʸ, t̥ʸs, g̱ʷ, d̥, n̥, r’, l̥, xʸ, m̥, t̥, x˙, r̥, g’, ð, br, rʷ, n̥, z̥, d̥, k̥, t̥s, s̥, x̥, x̌, l̥, ŋ̊, H, R, V, x̣˙, s’, l‘, g̱, hʷ, k˙’, kx, k̥x, L, t̥s, dð, tθ, t̥‘, pʸ, g‘, fʷ, t’θ, pf, p̥, Y’, mʸ, s̥, tʸ.

---

[5] In the case of Maidu the stops /b/, /d/, and /g/ are implosive, but since in no case in the data did an implosive series contrast with explosive sounds, these are listed simply as voiced, and the implosiveness is considered to be immaterial phonetic data.

[6] Hans Wolff, Osage I: Phonemes and Historical Phonology, IJAL.18.63–68 (1952).

TABLE I

This table contains all of the phonemes occurring in the sample with a frequency of occurrence of three percent or more.

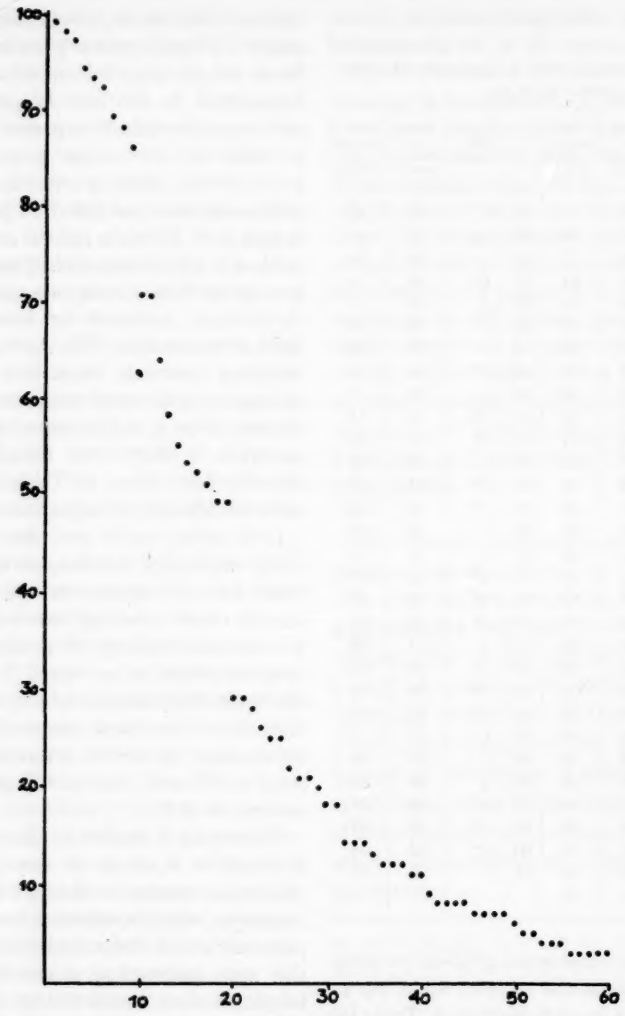
	Phoneme	Freq. of occurrence	Percentage of sample		Phoneme	Freq. of occurrence	Percentage of sample
1	k	173	99	32	f	25	14
2	t	172	98	33	dz	25	14
3	n	169	97	34	x̣	25	14
4	s	164	94	35	tɬ	24	14
5	m	163	93	36	y	23	13
6	p	161	92	37	θ	22	12
7	y	155	89	38	ž	21	12
8	h	153	88	39	k̓ʷ	21	12
9	w	151	86	40	t̓	18	11
10	tš	124	71	41	t'ɬ	18	11
11	l	123	71	42	lʸ	16	09
12	š	114	64	43	g	15	08
13	ts	103	58	44	gʷ	14	08
14	b	97	55	45	xʷ	14	08
15	x	93	53	46	t'	14	08
16	ʔ	92	52	47	k'	13	07
17	d	91	51	48	k'ʷ	13	07
18	g	86	49	49	dl	12	07
19	r	86	49	50	p'	12	07
20	k'	52	29	51	g̓	11	06
21	t'	52	29	52	t's	09	05
22	dž	51	28	53	kʷ	08	05
23	ɬ	46	26	54	t'š	07	04
24	p'	45	25	55	n'	07	04
25	ŋ	45	25	56	r̃	07	04
26	t's	39	22	57	x̣	06	03
27	k̓	37	21	58	k'	06	03
28	z	36	21	59	m'	06	03
29	t'š	35	20	60	g̓'	05	03
30	nʸ	32	18	61	l'	05	03
31	kʷ	32	18				

It is apparent that some symbols occurred in virtually all of the sample, while others occurred much less frequently (cf. Table I).

Since in many cases of scientific investigation a high frequency of occurrence of a given phenomena does not necessarily give that phenomena any significance, the data in Table I were plotted on cartesian coordinates (cf. Graph I). The assumption underlying this procedure is that a randomly selected set of data will form a smooth curve; changes in type of distribution, or breaks in the curve, would indicate significantly different types of phenomena so that these breaks may be considered as class boundaries. In the case of percentages of occurrence it would be expected that plotted in order of descending magnitude these points would form a smooth curve from somewhere near one hundred percent down to near zero. If this is not the case, then one method of empirically classifying consonants is to set up those consonants which form one distribution, bounded by breaks, as one class of consonants. This particular way of handling materials frequently brings out sharply a clustering of elements which is not obvious from a simple examination of the materials in chart form. Graph I contains the phonemes listed in Table I plotted in order of descending magnitude.

It is rather easily seen that the points, which represent the frequencies of occurrence for the consonants do not form a smooth curve starting near one hundred percent and trailing off gradually toward zero, as might be expected. Instead there are two rather sharp breaks in the distribution. The first break occurs between the consonants /w/, with a percentage of occurrence of 86, and /tš/, with a percentage of occurrence of 71.

This break is smaller in this sample than it would be if all of the descriptions were phonemic, because of the fact that in some languages what is recorded here as a /tš/ phoneme would undoubtedly be a cluster. If this were analyzed as a cluster in several languages, then the percentage of occurrence would be reduced below the 71 percent mark, thus increasing the size of the gap between /tš/ and /w/. This might seem somewhat irrelevant, since /l/ also has a percentage of occurrence of 71 percent. However, it has more relevance than it would at first seem, as good phonemicization would probably reduce the percentage of occurrence of /l/, because it is probably an allophone of /r/ in some instances. The



GRAPH I

The percentages of occurrence of the phonemes in Table I, plotted in order of descending magnitude.

important fact to consider here is not what the results might be if the method were more rigorous, but rather to point out the fact that the classification of consonants which results from this technique would be altered only in that the classes would be even more sharply defined if completely phonemic analyses had been available for all of the languages in the sample.

Consider first the group of consonants which occur above the break in the distribution; i.e. /k/, /t/, /n/, /s/, /m/, /p/, /y/, /h/, and /w/. The percentages of these consonants in order of descending magnitude are: 99, 98, 97, 94, 93, 92, 89, 88, 86. This descent appears to be linear, i.e. a single straight line could be drawn through the distribution so that the total distance from

all of the points of the distribution to that line would be less than if any other type of curve were so drawn. Linearity merely means that the difference between the percentage of occurrence of any two neighboring consonants is constant. Constancy here is relative, because in no sample of empirical data is a distribution exactly linear. However, the plot of a linear equation, i.e. an equation which represents a straight line in algebraic symbols, would give a very close fit, because in all three cases the drop between neighboring consonants is one percent, with a maximum drop of three percent. This is as near a perfect linear function as is likely to be found with this type of data.

One of the purposes of plotting distributions is to fit curves to them, and by extrapolating to predict the future course of the distribution. In simpler languages all the elements represented by a given group of data have something in common, i.e., they follow a certain curve. If the curve is extended beyond the particular class in question, it can be easily ascertained whether other data are a part of this distribution, and thus the class in question, or not. This is done by extending the curve, as predicted by its formula. If the data under investigation falls along the extended curve it belongs to the original group of data, because it has the specific qualities which cause the first group of phenomena to follow the distribution in question. If the second set of data does not follow the distribution, then it does not have those specific qualities.

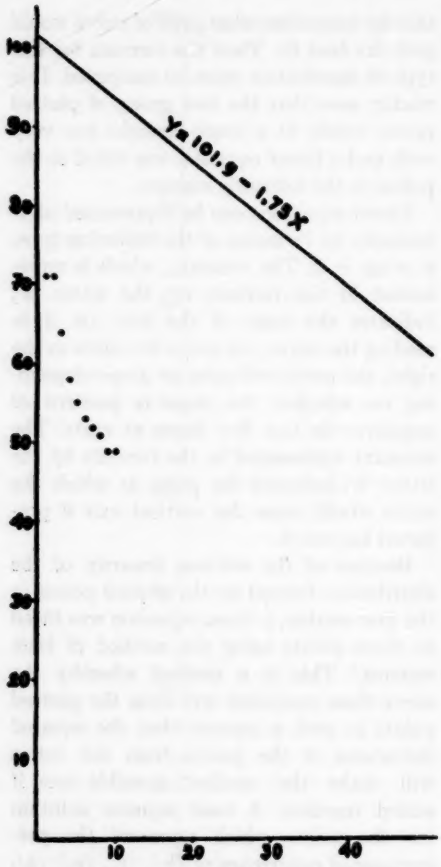
The above described technique will be applied to the data on Graph I. The purpose being to test whether the three sets of points which seem to cluster together form a single homogeneous distribution or three distinct distributions. This is done by computing the formula for the first distribution and projecting the curve. If the points in the second plot fit the curve, then there is one single homogeneous group of data represented by two samples, but if they do not fit, then the two plots represent independent groups.

The first step in the procedure is to ascertain by inspection what type of curve would give the best fit. Then the formula for this type of distribution must be computed. It is readily seen that the first group of plotted points would fit a single straight line very well, and a linear equation was fitted to the points in the following manner.

Linear equations can be represented algebraically by formulae of the following type: $y = ax + b$. The constant, which is represented in the formula by the letter 'a', indicates the slope of the line, i.e. if in reading the curve you move five units to the right, the curve will raise or drop—depending on whether the slope is positive or negative—5a (i.e. five times *a*) units. The constant represented in the formula by the letter 'b' indicates the point at which the curve would cross the vertical axis if projected beyond it.

Because of the obvious linearity of the distribution formed by the plotted points in the first section, a linear equation was fitted to these points using the method of least squares.[7] This is a method whereby the curve thus computed will miss the plotted points in such a manner that the squared deviations of the points from the curve will make the smallest possible sum if added together. A least squares solution for the points which represent the percentages of occurrence of /k/, /t/, /n/, /s/, /m/, /p/, /w/, /y/, and /h/ yielded the following equation; $y = 101.9 - 1.75x$. In order to predict the expected frequency for any given consonant, outside the first plot, simply count /k/ equals one, /t/ equals two, /n/ equals three, in order of descending magnitude of the consonant's frequency of occurrence, and substitute this number in the equation for x. Consider then the consonants /tš/ equals ten /l/ equals eleven, and /š/ equals twelve, which yield the expected percentages of occurrence of 84.4, 82.7, and 80 respectively. However, note that the observed percentages in Table I, for these are 71, 71, and 64 respectively. This disparity in

[7] Paul G. Hoel, Introduction to Mathematical Statistics, pp. 79–80 (1947).



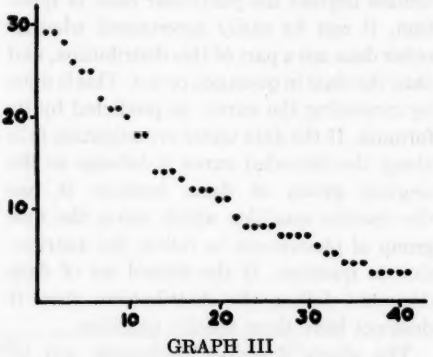
GRAPH II

These plotted points represent the second group of phonemes plotted in order of descending magnitude. The linear equation represents an extension of first distribution. Notice the deviation of the points from the line.

values means one of two things, either this is a temporary irregularity in the distribution, or these three values belong to a different distribution. This brings us to Graph II, which contains the percentage of occurrence of the consonants /tš/, /l/, /š/, /ts/, /b/, /x/, /ʔ/, /d/, /g/, and /r/ in that order (cf. Graph II).

Because of the shape of the distribution in Graph II, it would be very difficult to fit a curve by the method of least squares. By difficult is meant that the computation would be tedious. It is quite clear however from inspection of the plotted points that a linear equation would not represent the distribution realistically, simply because of the curvature in the plot. The linear equation, which fits the first group of points remarkably well, has been extended in Graph II to show that these plotted points in no way represent an extension of the first distribution. A curve was not fitted to the second plot, because the general shape is quite clear from inspection.

By constructing an imaginary line with the curvature indicated by the shape of the plot, the expected percentages of occurrence of the next three consonants in order, which are the first three of the third plot, are in the neighborhood of 47 percent. However, the observed percentages are 29, 29, and 28 for /k '/, /t '/, and /dž/ respectively. This brings us then to the third and final group of plotted points to be investigated.



GRAPH III

These points represent the distribution of the third and final group of consonants, i.e. the rare consonants, again plotted in order of descending magnitude.

p	t	k
	s	
m	n	
w	y	h

CHART II

Consonants typical of the sample as a whole

This distribution starts with /k '/ which has a percentage of occurrence of 29, and gradually slopes downward, approaching zero asymptotically, i.e. as the distribution is extended it steadily approaches zero as a limit, but never reaches it. Actually the curve does not approach zero as a limit, because we have a small finite sample. The real lower limit of the curve is .57 of one percent at which point the plot levels off parallel with the 'x' axis. If the size of the sample were increased the plot would gradually assume a curve which would approach the 'x' axis as a limit.

Looking again at the overall picture, there are clearly three distinct distributions. The first one is composed of the consonants /k/, /t/, /n/, /s/, /m/, /p/, /y/, /h/, and /w/ and can be interpreted as a typical consonant system, i.e. a model which was derived statistically and is typical of all languages in the sample. This system can also be said to be typical of all new world languages, to the extent that this sample is typical of the group as a whole.

To recapitulate briefly, the reasons for setting the above mentioned consonants off as in some way different from the rest of the distribution are: (1) the break in the distribution between /w/, and /tš/, and (2) the fact that the distributions above and below the break have quite different characteristics.

The drop in frequency of occurrence from /k/ to /w/, i.e. within group I is only 13 percent. The drop between neighboring consonants is one percent for all consonants except between /n/ and /s/, /p/ and /h/, and /h/ and /w/, which have a drop of 3, 3, and 2 percent respectively.

Compare this tight clustering with the gap of 15 percent between /w/ and /tš/, the last phoneme of group I and the first phoneme in group II. In addition to this consider the shape of the two distributions. The first group of plotted points forms a linear distribution, whereas the second would best be fit by a parabola or a hyperbola. These are alternate possibilities, and the way to test this would be to utilize the method of least squares[8] again, computing a polynomial (an equation with several powers of 'x' in this case) to see what type of equation resulted. However, this is a quite unnecessary task, as it is completely obvious that a linear equation will not fit these points. It is only necessary to prove non-linearity here to show independence, since the first distribution is distinctly linear.

The second break occurs between /r/ and /k '/, and the difference between these two consonants, i.e. between groups II and III, is 20 percent, while the maximum drop within group II (the groups are simply labeled I, II, and III respectively for convenience) is 7 percent, just over half the drop between groups I and II, and one third the drop between group II and III. This clearly shows the clustering within the three groups as opposed to the gaps between them.

The largest drop between consonants in the final group is 4 percent, and the curvature appears to be logarithmic in nature, but it is in no way an extension of the curve fitted to either groups I or II. This then is the empirical basis for setting up three groups of consonants. Group I is made up of consonants which appear in virtually all languages in the sample. Group III contains rare consonants, i.e. those which occur in a very small percentage of the sample. Group II is made of ten consonants which occur in just over half (from 49 to 71 percent) of the languages of the sample, but considerably below 90 percent, which represents the approximate percentage of occurrence for the typical consonants.

As stated earlier the consonants in group I were considered to constitute a model for New World languages, i.e. there is a very high probability that any new language found in North or South America will contain this group of consonants. Chart II below contains the consonantal system of the typical language in the New World.

[8] Idem.

NOTES AND REVIEWS

A Note on some Arawakan Words for *man*, etc.

Douglas Taylor

Although systematic comparative study of the large and widespread Arawakan linguistic stock must await the appearance of more and better descriptions of individual languages belonging to it, some insight may be gained in the meantime by considering such likely cognates, variants, derivatives and compounds as the data already published may reveal. The words for *woman*, *person* (*homo*), *man* (*vir*), *wife*, *husband*, *mother* (or rather such of them as I have been able to find recorded) constitute an especially interesting set, since they appear for the most part to have been complex and to contain or have contained a root morpheme in common. At the same time, phonetic reduction and reinforcement of various kinds have rendered reconstruction of this morpheme's prototype none too evident.

C. H. de Goeje, from whom the bulk of this material is taken, tells us that 'not all the various transcriptions for a given word have been copied, and complicated phonetic spelling has been simplified'; but in the main he follows the spelling of his sources, which often shows the influence of an author's national orthography (so, in the case of a German, *sch* and *tsch* for the sounds written *sh* and *ch* in English).[1] Since we cannot waste space by giving, as de Goeje does, references to the sources he consulted in the case of each language, the following conventions will be adopted here: s and ŝ represent, respectively, hissing and hushing sibilants; c and ĉ hissing and hushing sibilant affricates. When syllable initial, biiabial and palatal phonetic semivowels are shown as w and j respectively; elsewhere, u and i are employed both for syllabic and for non-syllabic sounds. y replaces both ö and ü and (syllabic) y of German authors, and represents what probably is a front rounded vowel in some languages, a back unrounded one in others. ñ is a palatal nasal, ĩ is a nasalized vowel (high front unrounded). Length is shown by gemination, as in Goajiro hiieer *woman* and haŝiĉĉi *man*. Other symbols have their usual phonetic values. No attempt at phonemicization has been made for the vast majority of the languages cited; nor is it possible with the data available to me.

The words for *woman*, *person*, *man*, *wife*, *husband*, *mother* are cited for 46 dialects (with some lacunae) as follows: 1 Island Carib, 2 Lokono, 3 Goajiro, 4 Paraujano, 5 Palicur, 6 Atorai, 7 Guinao, 8 Maipure, 9 Yavitero, 10 Baniva; 11 Bare, 12 Arequena, 13 Cariaia, 14 Carutana, 15 Catapolitani, 16 Siuci, 17 Yukuna, 18 Adzaneni, 19 Caouri, 20 Piapoco, 21 Mandauaca, 22 Jabaana, 23 Tariana, 24 Uainuma, 25 Mariate, 26 Cauixana, 27 Passé, 28 Uirina, 29 Araicu; 30 Maraua, 31 Cuniba, 32 Cujisenajeri, 33 Canamari, 34 Campa, 35 Piro; 36 Manao, 37 Mehinacu, 38 Jaulapiti, 39 Custenau, 40 Uaura, 41 Mojo, 42 Baure, 43 Saraveca, 44 Guana, 45 Terena, 46 Ariti (Paressi).

Woman: 1. hĩaru, 2. hiaro, 3. hiieer, 4. hñere, 5. ero, tino, 6. beanaro, 7. henau, 9. sariinaifemi, 12. inaru, 13. necu, henitaky, 14. inaru, 15. inaru, 16. inaru, 17. inanaru, 18. inaru, 20. inanai, 21. inaluĉa, 23. inaru, 24. inaru, 25. ynana, 28. inau, 30. runo, 36. ytunalo, 37. teneru, 38. tinau, 39. tineru, 40. teneru, 41. eseno, 42. eteno, 44. ceeno, 46. uilo.

Person: 1. ilaku, 2. loko, 3. wayuu, 7. kenawe, 9. enami, 10. enami, 11. heinari, 12. asinali, 14. acinari, 15. acinali, 17. asie, 18. naiki, aĉinali, 19. wenawika, 20. wenaika, naijei, 23. aciali, 26. cinani, 27. ŝimana, 29. etialo, 32. jehji, ŝeŝi, 35. ineri, 36. yrinaly ~ herenary, 37. erinau, 38. yrinau ~ erina, 39.

[1] See de Goeje, *The Arawak Language of Guiana* (Amsterdam, 1928), section 182, especially p. 219.

erina, enica, 40. eñau, 41. aĉane, 45. ŝáane, 46. aliti.

Man: 1. ejeri (/eiéri/), 2. wadili (/uadili), 3. haŝiĉĉi, 4. eije ~ eiĉ, 5. awaili, hiñeiri, 7, hanale, 9. enami, 10. enami, 11. heinari, 14. acinari, 15. acinali, 18. ciali, 20. aciali ~ asieri, 21. aĉinahe, 24. aĉijari, 28. atinare, 34. ĉirampari, ĉerari, 41. eh(o)iro, 42. ehiro, 43. eĉeena, 46. ina.

Wife: 1. iani, 2. iraito, 3. eeryĩ, 7. naune, 12. inu, 14. inu, 15. inu, 16. inu, 18. inu, 20. inu, 21. ju, 31. anendu, 32. hanando, 34. -hina, 35. anendu, 36. ero, 41. jeno, 46. ianitiu.

Husband: 1. iraiti, 2. ireti, 3. eeĉi, 7. jeme, 8. -mina, 10. mewale, 11. mi, 18. iniri, 19. niri, 20. niri, 21. niri, 22. imiri, imigi, 34. -hime, 35. aneri, 36. imiri, 41. ima, 44. uma, 45. ĩima, 46. ianini.

Mother: 1. -(u)kusuru, 2. -io ~ oro, 3. -eju, 8. ina, 18. ndoa, 20. atoa, 27. aiñu, 32. nato, 33. atu, 34. ina, inero ~ iniró, 35. ndo, 36. enakoni, 41. eeno, 42. eno, 44. eeno, 45. éeno.

In view of the deficiencies in the citations, both as regards lack of phonemicization and, especially, the many blanks it contains, the following remarks can offer only a tentative hypothesis which some readers may, it is hoped, help to confirm or to refute by supplying the missing data, and by furnishing new material from languages not cited here.

Several Arawakan languages have (productive) suffixes resembling, both formally and functionally, modern Lokono (True Arawak) -i ~ -li *male human* and -o ~ -ro *female human or non-human*; and it seems safe to say that such are or have been present in most of the words listed. Compare, for example, such pairs of words meaning *man*, *woman* as Catapolitani acinali, inaru; Uirina atinare, inau; Island Carib ijeri ~ ejeri (/eiéri/), hĩaru; and such others meaning *husband*, *wife* as Lokono ireti, iraito; Adzaneni iniri, inu; Piro aneri, anendu; Ariti ianini, ianitiu. Needless to say, these suffixes have in many cases become petrified; so that their one-time presence is revealed only by comparison with other forms in which they are still functional.

Ariti (Paressi) ina *man* and Maipure ina *mother* appear to be, despite the difference in meaning, cognate, simple lexical forms; and since in Campa we find both ina and inero ~ iniro for *mother*, we may assume that the first alternative is contained in the second, which shows two degrees of assimilation of the second vowel to the first. Similarly in the case of Piro ineri *person*, and in that of Adzaneni iniri *husband*; while such assimilation has been followed by loss of the first vowel in Caouri, Piapoco and Mandauaca niri *husband*. Cujisenajeri nato, Adzaneni ndoa, and Piro ndo, all meaning *mother*, show, respectively, loss of the first vowel without assimilation of the second, loss of the first vowel and shifting of the second to word final, and loss of both first and second vowels of the stem. Mojo, Baure, Guana and Terena e(e)no *mother* may plausibly be explained as resulting by regressive assimilation from *inau (cf. Uirina inau *woman*). Such a series among the words meaning *man* as: heinari (Bare), hiñeiri (Palicur), ejeri (Island Carib); or, among those meaning *woman*: inaru (many languages), hñere (Paraujano), hĩaru (Island Carib), hiaro (Lokono), must convince us that *(-)ina(-) has often lost its nasal component by way of palatalization; and we may suspect this to have happened in the case of Ariti iani-*spouse* and in that of Island Carib iani *wife*, with which compare Cauixana cinani *person*. Coalescence of morphemes is apt to follow —or to precede—such drastic phonetic reduction and/or assimilation; thus the stem of Lokono hiaro *woman* seems to occur in no other combination (unless -io *mother* < *-iño < *-inau), that of Lokono wadili (/uadili/) *man* probably being quite unrelated to it (cf. Lokono uadi- *long*, auadi- *go about, seek, beckon*).

In the case of about half the words for *woman*, and in that of most of those for *man* (some of which are also employed for *person*),

our hypothetical root *(-)ina(-) is preceded by what appears to be another morpheme (prefix? first member of a compound?), whose nature and function are obscure. Found both in words meaning *woman* and in others meaning *man* and/or *person* are various such forms which may probably be reconstructed as *at(i)-. Distinct from it, however, was another, *ar(i)-, found only in words meaning *person*; so compare Jaulapiti tinau, Custenau tineru and Mehinacu teneru, meaning *woman*, with yrinau ~ erina, erina, and erinau of the same languages, meaning *person*.

Of particular interest are: Caouri wenawika, Piapoco wenaika, Adzaneni naiki, all meaning *person*, since here we seem to have a reduplication of *winak (> *winak-winka > wenawika and further reductions), a form highly reminiscent of Proto-Mayan *winaq *person* (*homo*). Obviously, the rapprochement's significance would be enhanced if it could be shown that congeners of this form were widespread throughout the Arawakan stock, instead of being confined to a few languages. Other instances of reduplication, while not numerous, may be found in this list; so (though most probably unrelated to the foregoing) Cujisenajeri jehji and šeši, both meaning *person*, and partially reduplicated Yucuna inanaru, Piapoco inanai, Mariate ynana, all meaning *woman*. Moreover, Cujisenajeri hanando, Cuniba and Piro anendu, meaning *wife* are probably to be explained as having developed from *enan(a)ro; while Atorai beanaro *woman* might well have resulted from earlier *wenanaro by palatalization and loss of the first nasal, with strengthening of the initial bilabial.

Apart from this last example and the Caouri, Piapoco and Adzaneni words cited above, there is not much evidence for the loss of root-initial w; and it is at least clear that the w-sounds of wenawika *person* cannot have belonged to a prefix if this word is, as we assume, a reduplicated form. However, we may note the rounding and/or backing of the i of *(-)ina(-) in Mariate ynana (< *wina-na), Manao ytunalo (< *at-wina-lo), and Maraua runo (< *ar-wina-u), all meaning *woman*. Passé šimana *person* might plausibly be explained as a development of *(a)ti-wena by assimilation of w to the following nasal; and similarly Mojo ima, Guana uma, Terena šima, whence by combination with -ri Jabaana and Manao imiri, all meaning *husband*, from *wina through infection of n by w followed by loss of the latter root-initial. However, this last development is rendered unlikely by the fact that it has not taken place in the same languages' words for *wife* or *mother*.

There is still less trace of a root-final k; and there are no grounds for believing that the terminations of wenawika, wenaika, naiki were suffixal. But this is hardly surprising, since syllable-final consonants, especially stops, are rare or non-existent in many Arawakan languages; and in more than one of them there is evidence of their loss in historic times. So that if such a form as *winak were reduplicated, one would expect what one finds in attested wenawika, loss of one k and change of the other's position. Furthermore, if de Goeje should be wrong in regarding the k of Guinao kenawe (which he writes k-enaue) as a prefix, this word may well have resulted by metathesis from an earlier *wenake.

DOMINICA, BRITISH WEST INDIES

LANGUAGE IN RELATION TO A UNIFIED THEORY OF THE STRUCTURE OF HUMAN BEHAVIOR: Part II (Chapters 8–10). By Kenneth L. Pike. Glendale, California: Summer Institute of Linguistics, 1955. Pp. v + 85.

STANLEY NEWMAN

Like the last two or three chapters of Part I, published in 1954, the present volume deals primarily with language phenomena rather than with behavioral phenomena in their interrelated verbal and non-verbal aspects. Continuing the examination of minimum units (begun in Chapters 6 and 7 of

Part I), Chapter 8 discusses The Minimum Unit of the Manifestation Mode of the Behavioreme (1–40). In Pike's terminology this minimum unit is the acteme, appearing in non-verbal behavior as the kineme and in verbal behavior as the phoneme. The first half of this chapter treats the phoneme in terms of 'trimodal structuring,' a formulation which was established in Part I as a basic concept for the analysis of behavior. The FEATURE MODE of a phoneme is actualized in its physical components and serves to identify one phoneme in contrast to other phonemic units; here Pike presents his logic for rejecting the neutralization hypothesis (4–5). The MANIFESTATION MODE of a phoneme comprises its physical variants of differing types. In its DISTRIBUTION MODE a phoneme may be described by its potential for occurring at certain spots in phoneme sequences and within the more inclusive phonological units. The chapter then moves into a discussion of phonemic classes and systems, with their trimodal structure. The theoretical treatment is illustrated by an analysis of English phoneme classes, based largely on Bloomfield (24–30). The chapter ends with a summary presentation of phonetic theory and with a series of bibliographical comments on topics previously discussed.

Chapter 9, Higher-Layered Units of the Manifestation Mode of the Uttereme (41–70), is the first of three chapters which examine the more inclusive units in the hierarchy of linguistic form. Above the phoneme—the minimum emic unit of the manifestation mode of the uttereme, or verbal behavioreme—are such 'hyperphonemes' as the emic syllable, the emic stress group, the emic pause group, and other still larger phonological units. Interpreting the data collected by Stetson, Pike develops a distinction between the emic syllable, physiologically identified by a chest pulse, and the emic stress group, identified by an abdominal pulse. In Chapter 10, Higher-Layered Units of the Feature Mode of the Uttereme (71–74), Pike briefly discusses the hierarchy of units above the morpheme, such as the word, phrase, clause, sentence, etc. The total plan of Pike's work is to contain fourteen more chapters, whose titles are listed, following the table of contents, as being in preparation.

Compared to his approach in the preceding volume, Pike is on surer ground in the present publication. Instead of attempting to tackle the vast and amorphous area of behavioral theory, he confines himself largely to linguistics and, in particular, to phonological theory. The bibliographical comments appearing at the end of each chapter contain valuable historical summaries of many phonological concepts as well as critical evaluations of the theoretical positions taken by linguists of the past half-century, both American and European. His broad and intimate familiarity with linguistic theory provides a contrast to his much less certain grasp, displayed in Part I, of the literature concerned with behavioral theory.

But, in Pike's view, the range of linguistics itself is extensive and multidimensional. The phoneme, for example, is no monolithic entity. It is treated as a flexible and many-faceted unit with variable components, variable manifestations, and variable distributions; it forms part of a system which shows the same classes of variables; it is the minimum functioning unit in a hierarchy containing larger phonological units, which are also trimodally structured and form systems of their own at successive levels. Whether it be a phoneme or a syllabeme, the phonological unit exhibits differing manifestations in different styles of speech—in a song, or in a shout, in a whisper, or in a harsh voice (48); and these stylistic manifestations are structured in a manner which is linguistically significant. Consistent with this point of view, Pike's chief objection to the psycholinguists 'is not in reference to their general program or purpose, with which I am in hearty accord, but in reference to an exaggerated dichotomy which they might promulgate . . . between linguistics and psycholinguistics' (35). He regards the study of native speakers' reactions to the utterances

of each other, the observation of their attempts to write their language or to learn a foreign language, and presumably the use of experimental methods as proper tasks for the linguist.

Few linguists at the present time would object to Pike's aim in extending the boundaries of their science and in viewing linguistics against the perspective of human behavior. But objections can be raised to the way in which he believes this aim is to be pursued. On a practical level, the linguist is not trained to select the most useful methods nor oriented to formulate the most useful concepts for the observational and experimental approach to human behavior. These are methods and concepts which have received the concentrated attention of specialists in other disciplines—of biologists, psychologists, anthropologists, sociologists. And the linguist would be naive to assume that he can pretty well figure out these things by himself with ordinary common sense. On a theoretical level, it is equally naive to assume that the linguist is competent to construct a valid theoretical frame for the totality of human behavior merely by projecting his picture of linguistic structure on to a screen which looks as big as human behavior. Any such theoretical frame, to be genuinely all-inclusive, will come as the result of patient collaboration between the linguist and the specialists in other behavioral disciplines.

UNIVERSITY OF NEW MEXICO

LINGUISTIC SCIENCE AND THE TEACHING OF ENGLISH. Inglis Lectures in Secondary Education, 1954. By Henry Lee Smith, Jr. Cambridge: Harvard University Press, 1956. 61 pp.

GEORGE MOTHERWELL

This book is a publication of the Inglis Lecture addressed in 1954 to American educators, psychologists and anthropologists not trained in linguistics. Its implied purpose is to explain the educational significance of linguistic science and to appeal for increased interdisciplinary cooperation among scholars of these related fields—with a common goal of improved materials for the instruction of both teachers and students of English at the elementary and secondary levels. The relationship of language to culture is that proposed by Trager and Hall[1] and the Trager-Smith analysis[2] supplies the illustrations of English structure.

Much of Smith's material is exposition of concepts so apparent to the linguist that he rarely feels need to speak of them, and herein lies the value of the book. It will help to fill a need strongly felt by those concerned with linguistics in English and foreign language departments of American institutions: that for practical but undistorted exposition of basic concepts tacitly assumed in linguistic discussions but unfamiliar to the nonlinguist. The common confusion of language and writing (8, 14, 19) is one of the principle barriers not only to language learning but to mutual understanding between the linguist and the teacher of English to foreign and native students. The notions of one language's being more 'logical' than another (p. 10), 'superior' to another (p. 9), or 'easier' to learn than another (21), are beliefs which the linguist feels little need to consider but which probably exist among the majority of educators concerned with the teaching of English.

Smith begins with general statements. Language is considered as one of the ten basic systems of culture suggested by Trager and Hall. The linguistic scientist is a culturologist, primarily a cultural anthropol-

[1] Edward T. Hall, Jr., and George L. Trager, The Analysis of culture. [Pre-publication edition for criticism and discussion; reproduced for the authors with the assistance of ACLS] Washington, D.C. (1953).

[2] George L. Trager and Henry Lee Smith, Jr., An Outline of English structure. SIL: OP 3 (1951).

ogist, and linguistics is a branch of anthropology in this sense. Of a three-part system of culture consisting of language, kinesics and vocal qualifiers, language is the most important. Meaning is another aspect or combination of aspects of culture; only differential meaning is employed by the linguist and analysis of referential meaning leads us into consideration of culture as a whole in the anthropological sense. Language is systematically related to[3] but at the same time different from writing and the educator at the elementary and secondary school levels should be aware of this in order to provide more effective pedagogical materials. The confusion of language and writing in English (a result of language's being a FORMAL activity in the Trager-Hall schema[4]) exists on two levels: confusion of speech forms and writing, and confusion of styles of speaking and styles of writing. The former is to be dealt with at the elementary level. Then follows an illustration of the differences between language and writing using English vowels and stress phonemes with a few examples of 'parts of speech' definition based on criteria of stress and order (section II). The third and final section turns to the traditional concept of the 'sentence' and proceeds toward redefining this notion on the basis of juncture and intonation.

Smith is optimistic in tone and considers the problems of English language teaching from the viewpoints of both the educational psychologist and linguist. Comments are made upon the educator's lack of awareness of '. . . the strides that have been made by linguistic science in the last decades, and the extent to which this progress in the field of analyzing and describing the structure of languages can be applied to the very core of the language-arts program in the elementary grades and the education concerning language and carried on through language in the secondary school.' (4). He states that the culturologist and the psychologist are asking questions about the same phenomena although the psychologist is interested in the individual and the cultural anthropologist in the way that systems of culture are '*structured*, *interrelated*, and *transmitted*. . . . Both approaches are valid; indeed both approaches are essential, if vast areas of the utmost importance are to be systematically and fruitfully explored.' (6).

As for the application of linguistics as presented by Smith to the teaching of modern languages to non-natives, this reviewer feels that the materials developed under the auspices of the ACLS Intensive Language Program during World War II and at present speak for themselves. But it is well known that such large scale application has not been realized in the teaching of American English writing conventions in the standard English Composition course. The practical reasons for this are quite clear. First, linguistics has been concerned in the past largely with the phonemic and morphological levels of structure and the native speaker has little difficulty with these phenomena by the time he is of grade school age, whereas learners of English as a second language at any age must spend considerable time in mastering them. The second reason is that linguistic method was and is designed primarily for describing the structure of language rather than that of writing. Next is the preoccupation of linguists in the past with development and refinement of method more than with problems of pedagogy in this area; and finally, but perhaps no less important, is the strong conservative opposition found to the introduction of new methods found in English and language-arts programs in American universities.

But these are problems existing at the college level, and the point in connection

[3] Presumably the relationship revealed by graphemic analysis.

[4] Trager and Hall, p. 19. Such activities '. . . have to do with the establishing, learning, protection, and perpetuation of the unquestioned way to behave.' They (work, society, communication and the sexes) are learned, at least partially, out of awareness and taken for granted.

with Smith's book is that linguistics today will be of limited value to the college English teacher until it is applied to the materials of elementary and high school instruction in usage of spoken and written English. Even assuming extensive formal training in structural linguistics, a freshman English teacher's task will often consist not so much of building the student's knowledge of the language as of tearing down linguistic notions acquired in elementary and secondary schools which make it next to impossible to convey the importance of a distinction such as that between form and meaning. Therefore the crux of the matter is application, not at the college level, but before this—during the time when the native speaker's habits in usage of both form and meaning are congealing.

Although some of Smith's recommendations bring to mind a stately pleasure dome of the academic future with chairs in theoretical, experimental, applied and ethnolinguistics, some of his proposals have been partially realized in wide recognition of several new books which brighten the outlook for application. The most recent and notable work of this type is American English in Its Cultural Setting, written by two scholars who are not primarily linguists.[5] This book marks a long stride along the bridge from theory to practice which, it appears, will have to be seized before it is crossed. Smith's proposal that the relationship between language and '... the derived symbolic system that is writing ...' be recognized (15) is reflected in Chapter 21 of Lloyd and Warfel's book called Spelling: Phonemes Into Letters, in which appears a parallel list of English graphs and segmental phonemes of the Trager-Smith variety. Vocal qualifiers and kinesics are treated as essential qualities of communication,[6] and the bearing of the findings of geographic linguistics on standards of American speech and writing is shown. The 'culture' of the book's title does not indicate an examination of culture through systematic anthropological method and the list of phoneme-graph correspondences in Chapter 21 does not constitute graphemic analysis, but Smith's proposals are modest enough for us to consider this book as an example of a wish fulfilled.

Although an awareness on the part of educators of the nature of language in its relationship to culture '... would inevitably lead to a change in our teacher training programs ...' and '... could imply no less than a revolution in American education.' (7), the formidable practical barriers to be met are minimized by Smith's calling for no more than an awareness of the nature and functioning of language. He does not suggest that all pedagogues become linguists.

University of Southern California

[5] Donald J. Lloyd and Harry R. Warfel, American English in its cultural setting. Alfred A. Knopf, New York (1956).

[6] Lloyd and Warfel, pp. 61–62.

Letters and Announcements

1. Douglas Taylor (Dominica, British West Indies) reminds the editors that we did not include, in one of his recent articles (IJAL 22. 180–183) 'a footnote in reference to Dudley and Wyatt's vocabularies (1595) of Trinidadian Arawak.' These vocabularies, in Taylor's opinion, 'clinch the argument as to Lokono having been spoken in Trinidad at that time.'

2. Čestmír Loukotka (Praha-Strašnice, Czecho-Slovakia) submits the following bibliographical additions to Taylor's notes in IAJL 22. 180–183.

A. Insular Caribbean Languages.

Breton, Raymond

1664. Petit catechisme ou sommaire des trois premières parties de la doctrine chrestienne. Auxerre.

1665. Dictionnaire Caraïbe-français, meslé de quentinté de re-

marques historiques pour l'esclaircissement de la langue. Auxerre.
1666. Dictionnaire français-Caraïbe. Auxerre.
1667. Grammaire Caraïbe. Auxerre.
Davis, John
1666. The History of the Caribby Islands. London.
Neveu-Lemaire, M.
1921. Les Caraibes des Antilles. La Geographie t. XXXV 127–146. Paris. (Contains, on p. 144, a Dominica Carib vocabulary of 44 words.)
Ober, Frederick A.
Mss. Vocabulary of the Carib, Islands of Dominica and St. Vincent. In Bureau of American Ethnology, Washington.
Pelleprat, Pierre.
1655. Relation des missions de la Compagnie de Jésus dans les Isles et dans la terre ferme de l'Amérique meridionale. Paris.
Rochefort, César de.
1658. Histoire naturelle et morale des isles Antilles de l'Amérique. Rotterdam.

B. Language of the Black Caribs of Belize, etc.
Galindo, Juan.
1833. Notice of the Caribs in Central America. JRGS III. 290–291. London.
Henderson, Alexander.
1872. Grammar of the Karif Language as Spoken in the Bay of Honduras. Belize.
Lehmann, Walter.
1920. Zentral-Amerika. Die Sprachen Zentral-Amerikas. Berlin.
Membreño, Alberto.
1897. Hondureñismos. Tegucigalpa.
Sapper, Karl.
1897. Mittelamerikanische Caraiben. IAE X. 53–60. Leiden.
Stoll, Otto.
1884. Zur Ethnographie der Republik Guatemala. Zurich.

C. Sub-Taino of Cuba.
Culin, Steward.
1902. The Indians of Cuba. Philadelphia: Bull. of the Free Museum of Science and Arts. III. 185–226 and passim.

D. Eyeri of Trinidad.
Dudley, Roberto.
1661. Dell' arcano del mare. t. II. 33. Firenze. Reprinted in Penard, T. E., Remarks on an old vocabulary from Trinidad. De Westindische Gids. IX. 265–270, 1927, Amsterdam.

E. Nepuya of Trinidad.
Espinosa, Antonio Vazques de.
1948. Compendio y descripción de la Inias occidentales. SMC t. 108. 37. Washington.

3. A letter from Eva Sivertsen (Oslo, Norway) announces that The VIII International Congress of Linguists will be held in Oslo, Norway, August 5th to 9th, 1957. She adds: 'This is a closed congress, where the attendance, as far as voting members are concerned, is restricted to those who are invited by the Organizing Committee. However, we are open for suggestions as to linguists who should be invited. Those who would like to attend the Congress, but who have not yet received an invitation, are asked to apply to us for such an invitation.' Applications should be addressed to Cand. philol. Miss Eva Sivertsen, Secretary, Kirkeveien 98 A III, Oslo NV, Norway. Professor Alf Sommerfelt is President of the Congress.

Recent Publications

General Linguistics

Blixen, Olaf. La Linguística en el Uruguay: sus objetivos immediatos. Bol. de la Soc. de Antropologia del Uruguay 1.5–9, 1955.

Carroll, John B. (ed.). Language, Thought, and Reality. Selected Writings of Benjamin Lee Whorf. N. Y. and Boston: John Wiley and the Technology Press, M. I. T., 1956. xi + 278 pp.

Chalman, Seymour. Immediate Constituents and Expansion Analysis. Word 11.377–385, 1955.

Durand, Marguerite. De la perception des consonnes occlusive; questions de sonorité. Word 12.15–34, 1956.

Goodenough, Ward. Componential Analysis and the Study of Meaning. Lg 32. 195–216, 1956.

Greenberg, Joseph H. The Measurement of Linguistic Diversity. Lg 32.109–115, 1956.

Hoijer, Harry. Lexicostatistics: A Critique. Lg 32.49–60, 1956.

Izutsu, Toshihiko. Language and Magic. Studies in the Magical Function of Speech. Tokyo: Keio Inst. of Philological Studies, 1956. 140 pp. (In English)

Jakobson, Roman and Halle, Morris. Fundamentals of Language. (No. 1, Janua Linguarum.) 'S-Gravenhage (Holland): Mouton and Co., 1956. ix + 87 pp.

Lenneberg, Eric and Roberts, John M. The Language of Experience: A Study in Methodology. Indiana Univ. Pubs. in Anthropology and Linguistics, Mem. 13, 1956. 33 pp.

Malécot, André. Acoustic Cues for Nasal Consonants. Lg 32.274–284, 1956.

Malkiel, Yakov. The Uniqueness and Complexity of Etymological Solutions. Lingua 5.225–252, 1956.

Marchand, James W. Internal Reconstruction of Phonemic Split. Lg 32.245–253, 1956.

Paardekooper, P. C. Syntaxis Spraakkunst en Taalkunde. L. C. G. Malmberg (Den Bosch, Holland), 1955. xv + 311 pp. (In Dutch, with brief English summary)

Pickett, Velma. An Introduction to the Study of Grammatical Structure. Glendale, Calif.: Summer Inst. of Linguistics, 1956. 252 pp. (mimeog.).

Pike, Kenneth L. and Pike, Eunice V. Live Issues in Descriptive Linguistic Analysis. Glendale, Calif.: Summer Inst. of Linguistics, 1955. 23 pp.

Saporta, Sol. Morph, Morpheme, Archimorpheme. Word 12.9–14, 1956.

Smith, Henry Lee, Jr. Linguistic Science and the Teaching of English. Cambridge: Harvard Univ. Press, 1956. 61 pp. (Reviewed in this issue)

Townsend, Guillermo C. The Psychophonemic Method of Teaching to Read. America Indigena 16.123–132, 1956.

American Indian Languages

Bright William. Glottochronologic Counts of Hokaltecan Material. Lg 32.42–49, 1956.

Castro, Carlo Antonio. La lingüística en el Centro Coordinador Tzaltal Tsotzil. America Indigena 16.139–158, 1956.

Comas, Juan. La lengua vernácula y el bilingüismo en la educación. America Indigena 16.83–92, 1956.

Dozier, Edward P. Two Examples of Linguistic Acculturation: The Yaqui of Sonora and Arizona and the Tewa of New Mexico. Lg 32.146–157, 1956.

Haas, Mary R. Natchez and the Muskogean Languages. Lg 32.61–72, 1956.

Kluckhohn, Clyde. Some Navaho Value Terms in Behavioral Context. Lg 32. 140–145, 1956.

Lounsbury, Floyd. A Semantic Analysis of Pawnee Kinship Usage. Lg. 32.158–194, 1956.

McQuown, Norman A. Evidence for a Synthetic Trend in Totonacan. Lg 32.78–80, 1956.

Olmsted, David L. Palainihan and Shasta I: Labial Stops. Lg 32.73–77, 1956.

Read, William A. Four Indiano-Brazilian Lexical Notes. Romance Philology 9. 30–31, 1955.

Swadesh, Morris. Problems of Long-range Comparison in Penutian. Lg 32.3–16, 1956.

Taylor, Douglas M. An Additional Note on

the Consonantal System of Island Carib. Word 11.420–423, 1955.

Voegelin, C. F. Phonemicizing for Dialect Study (with reference to Hopi). Lg 32.116–135, 1956.

Publications Received

America Indigena. v. 16, nos. 2, 3, 1956.

Archivum Linguisticum. v. 7, fasc. 2, 1955.

American Speech. v. 31, no. 2, 1956.

Barton, Roy F. The Mythology of the Ifugao. Mem., American Folklore Soc., v. 46, 1955.

Boletin de la Sociedad de Antropologia del Uruguay. Año 1, no. 1, 1955. (Mimeog.)

Ethnohistory. v. 3, no. 1, 1956.

General Linguistics. v. 1, no. 3, 1955.

Japanese Journal of Ethnology. v. 19, nos. 3, 4, 1956.

Language. v. 32, nos. 1, 2, 1956.

Lingua. v. 5, no. 3, 1956.

Lord, Albert Bates (ed.). Slavic Folklore: A Symposium. American Folklore Soc., 1956. viii + 132 pp.

Luomala, Katharine. Voices on the Wind: Polynesian Myths and Chants. Honolulu: Bishop Mus. Press, 1955. 191 pp.

Muslim World. v. 46, no. 3, 1956.

Oceania. v. 26, nos. 3, 4, 1956.

PMLA. v. 51, nos. 2, 3, 3, 1956.

Saucier, Corinne L. Traditions de la Paroisse des Avoyelles en Louisiane. Mem., American Folklore Soc., v. 47, 1956. vii + 162 pp.

Word. v. 11, no. 3, 1955; v. 12, no. 1, 1956.

WILEY — BOOKS —

New insights into the communication sciences

ON HUMAN COMMUNICATION

A Review, a Survey, and a Criticism

By COLIN CHERRY, *Imperial College, University of London.* A fascinating series of papers which review recent trends in the communication sciences. It outlines the causes and the growth of modern interest in the relations between the diverse studies of communication such as linguistics, phonetics, information theory, experimental psychology, and semantics. The author explores the unification which exists in these areas, but at the same time he points out the difficulties of unification and the reasons which keep them separate. The book serves as an invaluable "source book" of references, citations, and definitions for all workers in the field of communication.

1957 **333 pages** **Illus.** **$6.75**

Major papers of a remarkable linguistic scientist

LANGUAGE, THOUGHT, AND REALITY

Selected Writings of Benjamin Lee Whorf

Edited and with an introduction by JOHN B. CARROLL, *Harvard University. Foreword by* STUART CHASE. Gathers together in a single convenient volume the principal contributions of one of the first popularizers of modern linguistics. It includes nearly all of Whorf's writings relating to the hypothesis that the structure of language influences thought processes and our perception of the world around us, plus a number of selections not published previously.

1956 **278 pages** **Illus.** **$7.00**

The past, present, and future of a fast-moving new field

MACHINE TRANSLATION OF LANGUAGES

Fourteen Essays, edited by WILLIAM N. LOCKE, *The Massachusetts Institute of Technology, and* A. DONALD BOOTH, *Birkbeck College Computational Laboratory, London.* The first book to survey mechanical translation. Seventeen top-ranking men have joined forces to present a well-documented story of MT's origins, a thorough treatment of the current state of the art, and a guide post to future lines of research.

1955 **243 pages** **$6.00**

Co-Published
by THE TECHNOLOGY PRESS
of The Massachusetts Institute of Technology
and JOHN WILEY & SONS, Inc.
440 Fourth Avenue New York 16, N. Y.

INDIANA UNIVERSITY PUBLICATIONS IN ANTHROPOLOGY AND LINGUISTICS

1. Paul Radin, Winnebago Hero Cycles: A Study in Aboriginal Literature (168 pp., 1948) $2.00
2. Paul Radin, The Culture of the Winnebago: As Described by Themselves (119 pp., 1949) $2.00
3. Paul Radin, The Origin Myth of the Medicine Rite: Three Versions. The Historical Origins of the Medicine Rite (78 pp., 1950) $2.00

4–5. Harold E. Driver and S. H. Riesenberg, Hoof Rattles and Girls' Puberty Rites in North and South America (33 pp., 1950); and Hilda J. Curry, Negative Painted Pottery of Angel Mounds Site and its Distribution in the New World (57 pp., 1950) $1.00

6. L. S. Freeland, Language of the Sierra Miwok (205 pp., 1951) $3.00
8. Claude Levi-Strauss, Roman Jakobson, C. F. Voegelin and Thomas A. Sebeok, Results of the Conference of Anthropologists and Linguists (86 pp., 1953) $1.00
9. Harold E. Driver, John M. Cooper, Paul Kirchhoff, Dorothy Rainier Libby, William C. Massey, and Leslie Spier, Indian Tribes of North America (Map + 30 pp., 1953) $1.00. Outline maps 11 x 17 inches may be purchased separately
11. Charles F. Hockett, A Manual of Phonology (246 pp., 1955) $3.50
12. Harold E. Driver, An Integration of Functional, Evolutionary, and Historical Theory by Means of Correlations; Bjarne Ulvestad, An Approach to Describing Usage of Variants in Language (59 pp., 1956) $1.50
13. Eric H. Lenneberg and John M. Roberts, The Language of Experience (33 pp., 1956) $1.00

PUBLICATIONS OF THE RESEARCH CENTER IN ANTHROPOLOGY, FOLKLORE, AND LINGUISTICS

1. Edward Sapir and Morris Swadesh, Native Accounts of Nootka Ethnography (457 pp., 1955) $5.00
2. Eeva K. Minn, Studies in Cheremis, Volume 4, Derivation (viii + 99 pp., 1956) $2.00
3. Robert N. Pehrson, The Bilateral Network of Social Relations in Konkama Lapp District (xiv + 128 pp., 1957) $2.50

Volumes of both the above series are available separately and are also distributed to IJAL subscribers. Orders for the IUPAL series may be sent to Mrs. Elsie Dosch, Business Manager, IJAL and IUPAL, Department of Anthropology, Indiana University, Bloomington, Ind. Orders for the PRC-AFL, to Thomas A. Sebeok, Director of Publications of the Research Center in Anthropology, Folklore, and Linguistics, Indiana University, Bloomington, Ind.

CONTENTS

A Tentative Typology of Etymological Studies.....*Yakov Malkiel* 1

Yuman Languages I: First Impressions.....*Werner Winter* 18

Phonemes of Ocaina (Huitoto).....*Arlene Agnew and Evelyn G. Pike* 24

The Frequency of Consonant Clusters in Chontal
Kathryn C. Keller and Sol Saporta 28

A Statistical Study of Consonants in New World Languages I: Introduction
Joe E. Pierce 36

NOTES AND REVIEWS

A Note on Some Arawakan Words for *man*, etc.....*Douglas Taylor* 46

Language in Relation to a Unified Theory of the Structure of Human Behavior (Pike)
Stanley Newman 48

Linguistic Science and the Teaching of English (Smith).....*George Motherwell* 50

Letters and Announcements..... 52

Recent Publications..... 53

Publications Received..... 55

COMPOSED AND PRINTED BY WAVERLY PRESS, INC., BALTIMORE, U.S.A.